ETSI TS 129 303 V8.2.0 (2009-06)

Technical Specification

Universal Mobile Telecommunications System (UMTS);
LTE;
Domain Name System Procedures;
Stage 3
(3GPP TS 29.303 version 8.2.0 Release 8)



Reference
RTS/TSGC-0429303v820

Keywords
LTE, UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2009.
All rights reserved.

DECT™, **PLUGTESTS**™, **UMTS**™, **TIPHON**™, the TIPHON logo and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.
3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.
LTE™ is a Trade Mark of ETSI currently being registered
for the benefit of its Members and of the 3GPP Organizational Partners.
GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Contents

Intellectual Property Rights 2
Foreword 2
Foreword 5
1 Scope 6
2 References 6
3 Definitions, symbols and abbreviations 7
3.1 Definitions 7
3.2 Abbreviations 7
4 General DNS Based Node Selection Description 7
4.1 Resource Records 7
4.1.1 A and AAAA 7
4.1.2 NAPTR 8
4.1.3 SRV 8
4.2 Selecting Domain Names 8
4.3 Identifying Nodes, Services and Protocols 8
4.3.1 IETF RFC 3958 Service and Protocol service names for 3GPP 8
4.3.2 Identification of canonical node names 9
4.3.3 Services from node names or other FQDN identifying a service 10
4.3.3.1 General 10
4.3.3.2 Procedure 10
4.3.3.2.1 S-NAPTR Procedure - General 10
4.3.3.2.2 S-NAPTR Procedure for a canonical node name 11
4.3.3.3 Services of a PGW from PGW node name (or collocated PGW/GGSN) 11
4.3.3.4 Services of a MME from MME node name (or GUTI) 12
4.3.3.5 Services of an SGSN from a P-TMSI 13
4.3.3.6 Services of a SGW from SGW canonical node name 14
5 Procedures for EPC Node Discovery and Selection 15
5.1 Procedures for Discovering and Selecting a PGW 15
5.1.1 Discovering a PGW for a 3GPP Access 15
5.1.1.1 General 15
5.1.1.2 Discovering a PGW or collocated PGW/GGSN for a 3GPP Access - S8/Gp roaming case existing PDN 16
5.1.1.3 Discovering a PGW or collocated PGW/GGSN for a 3GPP Access - S5/Gn intra-operator existing PDN 17
5.1.1.4 Discovering a PGW, collocated PGW/GGSN or GGSN for a 3GPP Access - S5/Gn intra-operator initial attach 18
5.1.2 Discovering a PGW for a non-3GPP Access with Network Based Mobility Management 19
5.1.2.1 Discovering a PGW for a non-3GPP Access – S2a/S2b initial attach for roaming and non-roaming 19
5.1.2.2 Discovering a PGW for a non-3GPP Access – S2a/S2b initial attach and chained S2a/S2b with GTP or PMIPv6 based S8 19
5.1.3 Discovering a PGW for a non-3GPP Access with DSMIPv6 19
5.1.3.1 Discovering a PGW for a non-3GPP Access – S2c initial attach 19
5.2 Procedures for Discovering and Selecting a SGW 20
5.2.1 General 20
5.2.2 SGW Selection during TAU or RAU with SGW change - 3GPP roaming case 20
5.2.3 SGW Selection during TAU or RAU with SGW change - non-roaming case 21
5.2.4 SGW Selection during non-3GPP handover to 3GPP access 22
5.3 Procedures for Discovering and Selecting a PGW and SGW 22
5.4 Procedures for Discovering and Selecting an MME 23
5.5 Procedures for Discovering and Selecting an SGSN 24
5.5.1 General 24
5.5.2 SGSN initial target selection based on RAI 24

5.5.3 SGSN initial target selection based on RNC-ID (UTRAN target) 25

Annex A (Informative): Examples 27
A.1 Introduction 27
A.2 Preconditions 27
A.3 Colocated Simple LTE Example 27
A.3.1 Network description 27
A.4.2 Master file for "Collocated Simple LTE Example" 28
A.4.3 SOA and NS records 28
A.4.4 MME file for "Collocated Simple LTE Example" 29
A.4.5 APN file for "Collocated Simple LTE Example" 31
A.4.6 PGW/SGW node file for "Collocated Simple LTE Example" 31
A.4.7 TAI/TAC file for "Collocated Simple LTE Example" 33
A.4.8 MME lookup based on GUTI for "Simple LTE Example" 35
A.4.9 APN lookup for "Simple LTE Example" (i.e. PGW candidate list) 36
A.4.10 TAI lookup for "Simple LTE Example" (i.e. SGW candidate list) 38
A.4.11 Finding the collocated SGW and PGW together 40
A.4.12 S11 lookup by SGW canonical node name 41
A.4.12 TAI lookup for "Colocation Simple LTE Example" (i.e. MME candidate list) 42

Annex B (Normative): DNS procedures clarifications 44
B.1 DNS RFC procedures general clarifications 44
B.2 DNS procedures 3GPP clarifications on S-NAPTR 44

Annex C (Informative): DNS Pseudo-Code 46
C.1 S-NAPTR procedure base pseudo-code 46
C.2 S-NAPTR procedure - no topon 47
C.3 S-NAPTR procedure candidate list 48
C.4 S-NAPTR procedure pseudo-code with topon 49

Annex D (Informative): Change history 52
History 53

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
 - 1 presented to TSG for information;
 - 2 presented to TSG for approval;
 - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document describes Domain Name System (DNS) Procedures for the Evolved Packet System. This document covers the Evolved Packet Core gateway node selection using DNS (e.g. SGW and PGW nodes) excluding all User Equipment (UE) initiated DNS-based discovery and selection procedures.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] IETF RFC 1034:"DOMAIN NAMES - CONCEPTS AND FACILITIES".

[3] IETF RFC 1035:"DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION".

[4] 3GPP TS 23.003: "Numbering, addressing and identification".

[5] GSMA PRD IR.67 – "DNS Guidelines for Operators" Version 2.1.0.

[6] IETF RFC 3596: "DNS Extensions to Support IP Version 6".

[7] IETF RFC 3403: " Dynamic Delegation Discovery System (DDDS) Part Three: The Domain Name System (DNS) Database".

[8] IETF RFC 2782: "A DNS RR for specifying the location of services (DNS SRV)".

[9] IETF RFC 3958: "Domain-Based Application Service Location Using SRV RRs and the Dynamic Delegation Discovery Service (DDDS)".

[10] IETF RFC 3401: "Dynamic Delegation Discovery System (DDDS) Part One: The Comprehensive DDDS".

[11] 3GPP TS 23.401: "GPRS enhancements for E-UTRAN access ".

[12] 3GPP TS 25.413: "UTRAN Iu interface RANAP signalling".

[13] IETF RFC 2671: "Extension Mechanisms for DNS (EDNS0)".

[14] IETF RFC 3402: "Dynamic Delegation Discovery System (DDDS) Part Two: The Algorithm".

[15] IETF RFC 2308: "Negative Caching of DNS Queries (DNS NCACHE)".

[16] IETF RFC 3330: "Special Use IPv4 Addresses".

[17] IETF RFC 3849: "IPv6 Address Prefix Reserved for Documentation".

[18] 3GPP TS 23.060: "General Packet Radio Service (GPRS); Service Description; Stage 2".

3 Definitions, symbols and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

Domain Name System as defined in IETF RFC 1034 [2], IETF RFC 1035[3], and as used in 3GPP in 3GPP TS 23.003 [4] and GSMA PRD IR.67 [5]

The phrase "operators shall provision" in this document is intended to convey what is required to provision in DNS to provide DNS based selection for the corresponding function documented here. If there is a non-DNS procedure in an operator's network for that function then there is no functional requirement for the operator to provision such DNS records.

The term "S4-SGSN" refers to a Release-8 SGSN that has at least one set of S4/S3/S16 interfaces enabled.

The term "Release 8 SGSN supporting only Gn/Gp" refers to a Release 8 or later SGSN that either explicitly does not support S4 interfaces or all S4/S3/S16 interfaces are disabled due to operator policy. Such a node cannot use an SGW but can use a collocated PGW/GGSN. See 3GPP TS 23.401 [11] Annex D for use cases.

The term "Release-8 SGSN" applies to either case.

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905[1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

DNS	Domain Name System
DDDS	Dynamic Delegation Discovery Service
FQDN	Fully Qualified Domain Name
GUTI	Globally Unique Temporary Identity
PGW	PDN Gateway
SGW	Serving Gateway
TAI	Tracking Area Identity
TAU	Tracking Area Update

4 General DNS Based Node Selection Description

4.1 Resource Records

4.1.1 A and AAAA

The A resource record is used to define IPv4 host address corresponding to fully qualified name of the host as defined in IETF RFC 1035 [3]. The AAAA resource record is used to define IPv6 host address corresponding to fully qualified name of the host as defined in IETF RFC 3596 [6].

It should be noted that in DNS A or AAAA record names, in general, represent a host and its "equivalent" interface. Host names, in general, cannot be used as node names. A node may need to have more than one host name for the simple reason that it can have multiple interfaces for different purposes.

4.1.2 NAPTR

The NAPTR resource record is defined in IETF RFC 3403 [7] and is a powerful tool that allows DNS to be used to lookup services for a wide variety of resource names, which are not in domain name syntax. NAPTR would be used by a client program to rewrite a string into a domain name. The rewrite process is controlled by flags that provide information on how to communicate with the host at the domain name that was the result of the rewrite. If DNS returns multiple NAPTR resource records those can be prioritized using embedded order and preference values defined by the DNS administrator.

The S-NAPTR procedure i.e., the "Straightforward-NAPTR" procedure, is defined in IETF RFC 3958 [9] and describes a Dynamic Delegation Discovery System (DDDS) [10] application procedures on how to resolve a domain name, application service name, and application protocol dynamically to target server and port by using both NAPTR and SRV (see IETF RFC 2782 [8]) resource records. The S-NAPTR also simplifies the use of NAPTR by limiting the NAPTR flags only to "a", "s" and "". Furthermore, only NAPTR "replacement" expressions are allowed, not "regular expressions", during the rewrite process. The changes compared to IETF RFC 3403 [7] NAPTR usage are procedural and are limited only to the resolver. The S-NAPTR use of the NAPTR resource record is exactly the same as defined in IETF RFC 3403 [7] from the DNS server and DNS infrastructure point of view. Additional information on S-NAPTR usage is provided in Annex B and Annex C.

The NAPTR resource record flags "s" and "" allow another layer of indirection in the DNS configuration. The "" flag causes the S-NAPTR procedure to query for new NAPTR resource records from the DNS infrastructure. The "s" flag causes the S-NAPTR procedure to query for an intermediary SRV resource record pointing to A/AAAA resource records. This additional query provides a selection mechanism by which the operator is able to assign different weights to different A/AAAA resource records while larger weights are given a proportionately higher probability of being selected. A DNS server might provide the A/AAAA records together with the SRV resource records as per IETF RFC 2782 [7]. The length of the NAPTR resource record indirection chain enabled using the "" flag is unbounded and may lead to a deep chaining of resource records over time in the DNS configuration. Additional layer of indirection and possible deep chaining both grows the DNS configuration significantly in size and complexity, and also makes the configuration prone to hard to trace errors. The use of NAPTR resource record "" flag pointing to other NAPTR resource records with flag "" is strongly discouraged. Specifically, NAPTR resource flag "" should only be provisioned to point to terminal NAPTR records (i.e., flag "a" or flag "s"). Generally, the use of flag "a" or of flag "s" is encouraged.

4.1.3 SRV

The SRV resource record is defined in IETF RFC 2782 [8] and allows DNS administrators to use pool of servers for a single domain with static load balancing to each server, to move services from host to host, and to designate some hosts as primary servers for a service from a pool of hosts. A resolver can ask for a specific service/protocol combination for a specific domain name and get back a Fully Qualified Domain Names (FQDN) of any available servers.

4.2 Selecting Domain Names

When using the S-NAPTR procedure under the DDDS framework, it becomes essential which domain name gets used for querying the actual NAPTR records. In the S-NAPTR procedure, the Application-Unique String used by the DDDS algorithm is the starting domain name for which the information of the services, protocols and actual canonical node names are sought. Related to the Application-Unique String, the First well-Known Rule of the DDDS algorithm in the S-NAPTR procedure outputs the same domain name that constitutes the Application-Unique String. For each node type in EPC that can be queried for information using the S-NAPTR procedure, the authoritative DNS server for the given domain should be provisioned with unique domain name for each EPC node or other identifier that is explicitly specified by a procedure in this specification (for example one based on APN, TAI,GUTI, etc) and corresponding NAPTR records. The authoritative DNS server for a given domain shall provision at least the EPC node names that may be exposed to the inter-operator roaming interfaces.

4.3 Identifying Nodes, Services and Protocols

4.3.1 IETF RFC 3958 Service and Protocol service names for 3GPP

Service and protocol service names for the S-NAPTR procedure shall be used in accordance with 3GPP TS 23.003 [4], subclause 19.4.3.

4.3.2 Identification of canonical node names

There are many use cases where it is desirable to select a collocated node in preference to a non-collocated node, or a topologically closer (with respect to the network topology) node in preference to a less topologically closer node. To easily do this action a "canonical" node name shall be employed so that the "canonical" node names from two or more sets of records can be compared to see if nodes are actually the same nodes, or topologically closer nodes.

In DNS neither A or AAAA host names, in general, represent a node name, but rather a set of "equivalent" interfaces. A node may need to have more than one host name for the simple reason that it can have different interfaces for different purposes. For example, a node can have a set of roaming interfaces on a completely different network than the internal network due to security needs. Hence, there are always situations where multiple A/AAAA record sets must exist that implies multiple distinct host names. Therefore, host names, in general, cannot be used as node names.

Instead of creating new DNS records to map a host name to a node name this specification defines how host names shall be constructed and used in S-NAPTR procedure within 3GPP EPC.

The host names shall have form:

<"topon" | "topoff"> . <single-label-interface-name> . <canonical-node-name>

Where the first label is "topon" or "topoff" to indicate whether or not collocated and topologically close node selection shall be preferred, "single-label-interface-name" is a single label used to name a specific interface on a node (e.g. Eth-0, S8, vip, board3), "canonical-node-name" is a the canonical node name of a specific node. A node shall have exactly one canonical node name so a host name always includes the unique canonical node name of the node.Hence, when comparing the host name FQDNs to find out whether the nodes are actually the same, the first two labels of the host name FQDN shall be ignored.

NOTE 1: The canonical node name is not related to canonical name in the CNAME DNS record.

When using host names with "topon" as the first label the canonical node names of nodes shall be hierarchically structured to allow an operator to reflect the topological closeness of two nodes by naming the nodes with canonical node names sharing a common suffix domain name. The number of labels in the common suffix shall represent how close the operator considers them during node selection. The higher the number of labels in the common suffix is, the closer the nodes are. In other words, two topologically closest nodes are those with the longest matching suffix in their respective canonical node names.

The following list contains examples of domain names where canonical node names are in bold:

topon.Eth-0.**gw32.california.west. example.com**

topon.S8.**gw32.california.west. example.com**

topon.vip.**sgw3.oregon.west. example.com m**

topon.board3.**pgw1.cluster1.net27. example.net**

topon.S5.**gw4.cluster1.net27. example.net**

topon.board3.**pgw1.cluster2.net27. example.net**

In the examples above, "Eth-0.gw32.california.west.example.com " and "S8.gw32.california.west. example.com " are two different interfaces on the same node, "gw32.california.west. example.com ". On the other hand, "gw4.cluster1.net27.example.net" is topologically closer to "pgw1.cluster1.net27. example.net" (they are both connected to the "cluster1.net27. example.net" subnetwork) than to "pgw1.cluster2.net27. example.net" (only connected to the wider "net.27. example.net" subnetwork.)

Interface names and node names do NOT identify a function in the procedures here. The interface is part of the natural hierarchy within a node and the host name is already returned with the existing DNS records. The approach of identifying a canonical node name from a host name is believed to be simpler and more logical to maintain than creating additional DNS records simply to return a node name.

The topologically aware naming restriction (i.e. the format above using "topon" or "topoff") shall be placed only on all targets/replacements pointing logically to a A/AAAA record sets from the S-NAPTR procedure. These targets/replacements are denoted host names here (following the normal DNS terminology unless a CNAME is used to

point to the actual A or AAAA record). This restriction shall NOT apply to any other DNS records the operator may be using.

Specifically, a NAPTR with flag "a" will have a replacement target pointing to the A/AAAA record directly, thus the topologically aware naming restriction on the host name applies to the replacement in the NAPTR record with a flag "a". For the NAPTR flag "s" case the topologically aware naming restriction on the host name applies to the target in the SRV record, and not the NAPTR record replacement. For the NAPTR empty flag "" case the topologically aware naming does not apply any restriction since this is not a host name. Other flags are not used in S-NAPTR.

During DNS provisioning for the S-NAPTR procedure the operator is free to add another layer of indirection using a CNAME record (see section 6.6.2 of IETF RFC 1034 [2]).

While operators shall provision host names in DNS according to the above rules, it is still possible that the host name might be incorrectly configured (i.e. not conforming to the above format with first label of "topon" or "topoff"). For such misconfigured records implementations shall treat the misconfigured host name as valid within the S-NAPTR procedure where that host name was found, but may favor correctly configured records. Misconfigured host name for topological matching and colocation checks shall be treated as if the misconfigured host name had the label "topoff." prepended. Operators shall not depend on this behavior.

The order of using DNS records to contact a node is based on following ordering. When collocation of a pair of node types is explicitly stated as applicable in a procedure, collocation of the nodes shall have high est importance . When topological matching is explicitly stated as applicable in a procedure then topological matching with "topon", is of second highest importance. Then finally the ordering obtained by the S-NAPTR output . When collocation and topological ordering applies, collocated sets of nodes have highest importance, then sets of nodes with "topon" and the most labels in their common suffix, then sets of nodes with "topon" and second most number of labels and so on until we reach non-colocated nodes with "topoff" and within sets with same colocation or topological order then S-NAPTR order of one of the two nodes is used (specified in the specific procedure). Additional informative clarifications on how S-NAPTR is employed in the context of 3GPP EPC node usage and specifically how topological matching using the "topon" label interacts with S-NAPTR ordering is provided in Annex C.4.

4.3.3 Services from node names or other FQDN identifying a service

4.3.3.1 General

There are potential use cases where a node has a logical name of a peer or other FQDN identifier for a service but does not have the protocols it supports. The NAPTR record for any of the services can be provisioned at the nodes logical name or other FQDN identifier for a service. The node logical name or other FQDN identifier for a service is equal to the domain name under which NAPTR records are provisioned. This allows any core network node to discover the available services based on node"s logical name or other FQDN identifier for a service.

4.3.3.2 Procedure

4.3.3.2.1 S-NAPTR Procedure - General

These procedures are employed when any core network node has the FQDN of an entity and needs to find one or more services at that entity.

NOTE: There are three likely sources of the entity name. O&M provisioned, 3GPP specified based on some other identifier of a service (such as GUTI, TAC, IMSI, MISDN etc.), or the canonical node name obtained from a previous S-NAPTR procedure. Note that a node can have more than one name, i.e. an alias, but there shall be only one canonical node name for a node. (CNAME records are a way to create aliases for the canonical node name or any other FQDN as per IETF RFC 1034 [2])

The S-NAPTR procedure requires that DNS NAPTR records shall be consistently provisioned as described in IETF RFC 3958 [9]. This means a NAPTR record for each protocol using "a" flag and the service field populated with the service and proto value may be provisioned. If a more sophisticated load balancing or non-standard ports are desired, NAPTR with "s" flag for each protocol and the corresponding SRV records with relative weighting for each interface need to be provisioned. NAPTR records with empty "" flag records may also be used.

A DNS resolver that intends to use the S-NAPTR procedure shall use the FQDN of the node or a specified FQDN identifier of a service as the Application-Unique String. If all protocols are desired, then the resolver simply runs the S-NAPTR procedure as if all protocols match.

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This list can be obtained one host name at a time, in a procedure similar to Annex C.2, or a complete ordered list of all nodes, in a procedure similar to Annex C.3. Such a complete list obtained from an S-NAPTR procedure is referred here as a candidate list.

NOTE: The candidate list is valid for at most for finite period of time due to DNS time to live and order of the records can change due to statistical selection. Operators should provision records with reasonable time to live values.

4.3.3.2.2 S-NAPTR Procedure for a canonical node name

One very important special case is S-NAPTR based on a node's canonical node name.

Operators shall provision NAPTR records for all enabled interfaces of a node that are explicitly listed in subclause 19.4.3 of 3GPP TS 23.003 [4] at the FQDN of the node's unique canonical node name.

NOTE 1: Exception to this rule: The NAPTR records for x-3gpp-mme: x- s1-mme and x-3gpp-sgw:x-s1-u from subclause 19.4.3 of 3GPP TS 23.003 [4] are to be considered entirely optional in this release of 3GPP.

With such provisioning by an operator a DNS resolver can at any time use the S-NAPTR procedure with a valid canonical node name to find all interfaces and protocols supported by that node that are listed in subclause 19.4.3 of 3GPP TS 23.003 [4]. This includes interfaces of co-located functions that might not be easily discovered by other means.

NOTE 2: The remaining subclauses within subclause 4.3.3 cover cases where the FQDN used for the Application-Unique String is known to correspond to the resource of exactly one node and are relatively simple. These include the cases where an explicit identifier is defined by 3GPP for S-NAPTR lookup and examples of S-NAPTR based on canonical node names (the later is not intended to be an exhaustive list of examples). More complicated cases that cover identifiers for multiple nodes are covered in other subclauses.

4.3.3.3 Services of a PGW from PGW node name (or collocated PGW/GGSN)

An UE with both a 3GPP access capability and non-3GPP access capability can roam in and out of the 3GPP network while maintaining the same PDN connection.

To support roaming to or from a non-3GPP network the HSS (or AAA) server can have a FQDN of a particular PGW or collocated PGW/GGSN node. One reason for using a FQDN instead of an IP address is that a PGW can be multihomed (i.e. more than one IP address). Another possible use case is when the PGW interface needs to be changed between PMIP and GTP. Even if each interface type only uses one IP address, the different interfaces can still use different IP addresses. For example, roaming and non-roaming interfaces are likely to be separated from each other using firewall or other mechanisms. Another possible use case is when the Home Agent (HA) functionality of a particular PGW needs to be discovered, e.g., during the HA reallocation procedure.

If the PGW node name employed by the operator is the PGW canonical node then see sub-clause 4.3.3.2 for provisioning.

If the PGW node name employed by the operator is not the PGW canonical node name then the operators shall provision at least the NAPTR records for S5 and S8 interfaces of a PGW node (see subclause 19.4.3 of 3GPP TS 23.003 [4]) at that PGW node name. If the GGSN function is co-located then the NAPTR records for the Gn/Gp interfaces of the collocated PGW/GGSN shall be included as well. Only interfaces that exist and are allowed by the operator policy need be included.

NOTE 1: The PGW node name in HSS/AAA may or may not be the PGW's canonical node name if the PGW's FQDN was placed in the HSS/AAA from a non-3GPP source.

To resolve the allowed PMIPv6 interfaces the S-NAPTR procedure shall be used with the "Service Parameters" of

"x-3gpp-pgw:x-s5-pmip" , "x-3gpp-pgw:x-s8-pmip"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the FQDN of the PGW or collocated PGW/GGSN node.

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This list can be obtained one host name at a time, in a procedure similar to Annex C.2, or a complete ordered list of all nodes, in a procedure similar to Annex C.3. Such a complete list obtained from an S-NAPTR procedure is referred here as a candidate list.

For a more explicit example, an operator might provision a PGW name at:

gw1.pgw.nodes.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

Similarly for the GTPv2 interfaces the S-NAPTR procedure shall use "Service Parameters" of

"x-3gpp-pgw:x-s5-gtp" , "x-3gpp-pgw:x-s8-gtp"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the FQDN of the specific PGW or collocated PGW/GGSN node.

Similarly for the GTPv1 Gn/Gp interfaces the S-NAPTR procedure shall use "Service Parameters" of

"x-3gpp-pgw:x-gn" , "x-3gpp-pgw:x-gp" "x-3gpp-ggsn:x-gn" , "x-3gpp-ggsn:x-gp"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the FQDN of the specific PGW , collocated PGW/GGSN node. The "Service Parameters" of "x-3gpp-pgw:x-gn", "x-3gpp-pgw:x-gp" represent collocated PGW/GGSN nodes and "x-3gpp-ggsn:x-gn", "x-3gpp-ggsn:x-gp" represent a GGSN that does not have a PGW co-located.

Similarly for the Home Agent functionality of a PGW the S-NAPTR procedure shall use "Service Parameters" of

"x-3gpp-pgw:x-s2c-dsmip"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the FQDN of the specific PGW.

It is also possible for the DNS resolver to leave "Service Parameters" unspecified in the S-NAPTR procedure in order to identify all interfaces for all supported services and protocols.

NOTE 2: The services based on S-NAPTR at the canonical node name of the PGW might return services other than ones starting with x-3gpp-pgw or x-3gpp-ggsn since the PGW node can have a co-located SGW function or other functions.

4.3.3.4 Services of a MME from MME node name (or GUTI)

There are procedures where the old MME must be contacted by the new MME or S4 SGSN(a Release-8 SGSN supporting only Gn/Gp may also optionally use this procedure). The primary use case is context transfer.

The 3GPP defined MME node FQDN shall be constructed as defined in subclause 19.4.2.4 of 3GPP TS 23.003 [4] where the needed data can be obtained from the UE's old GUTI (or mapped from old P-TMSI and old RAI see subclause 4.3.3.5 for more details). The 3GPP defined MME node FQDN is either the canonical node name itself or an alias of the MME's canonical node name (the operator is free to choose the canonical node name).

If the MME node name employed by the operator is the 3GPP defined MME node FQDN then see sub-clause 4.3.3.2 for provisioning.

If the MME node name employed by the operator is the 3GPP defined MME node FQDN the operator shall provision NAPTR records under the 3GPP defined MME node FQDN for at least "x-3gpp-mme:x-s10", "x-3gpp-mme:x-s3" if S3/S4 GERAN/UTRAN is supported, and "x-3gpp-mme:x-s16" if Gn/Gp is supported.

So, for example, for an MME to find all S10 interfaces of an MME based on the old GUTI the S-NAPTR procedure shall be prefixed with "Service Parameters" of

"x-3gpp-mme:x-s10"

and set the Application-Unique String to the FQDN as defined in subclause 19.4.2.4 of 3GPP TS 23.003 [4], with the initial query targeted at 3GPP defined MME node FQDN.

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This list can be obtained one host name at a time, in a procedure similar to Annex C.2, or a complete ordered list of all nodes, in a procedure similar to Annex C.3. Such a complete list obtained from an S-NAPTR procedure is referred here as a candidate list.

Similarly, for an S4 -SGSN to find all S3 interfaces of an MME based on the old GUTI it would use a "Service Parameter" of "x-3gpp-mme:x-s3".

Similarly, for a Release 8 Gn/Gp-SGSN to find all Gn and Gp interfaces of an MME based on the old GUTI it would use a "Service Parameter" of " x-3gpp-mme:x-gn", " x-3gpp-mme:x-gp".

To find all MME related services of an MME based on the MME's canonical node name the S-NAPTR procedure shall be prefixed with "Service Parameters" of

"x-3gpp-mme:x-s10", "x-3gpp-mme:x-s3", "x-3gpp-mme:x-s16", "x-3gpp-mme:x-s11", "x-3gpp-mme:x-gn", "x-3gpp-mme:x-gp",

and set the Application-Unique String to the MME's canonical node name.

It is also possible for the DNS resolver to use only the interfaces it is interested in or leave "Service Parameters" unspecified in the S-NAPTR procedure in order to identify all interfaces for all supported services and protocols.

NOTE 1: The services based on MME canonical node name can return services not starting with x-3gpp-mme since the MME node might, for example, have a co-located SGSN.

4.3.3.5 Services of an SGSN from a P-TMSI

There are procedures where the source SGSN must be contacted by the target MME or a target S4-SGSN. A Release 8 SGSN supporting only Gn/Gp may also optionally use this procedure.

During a mobility procedure towards a new core network node, a UE served by a SGSN has a previously assigned P-TMSI by the source SGSN. A pre-Release-8 UE will provide the P-TMSI. A Release-8 UE will map P-TMSI to a derived GUTI using the procedure in 3GPP TS 23.003 [4] sub-clause 2.8.2and referred to in Annex H of 3GPP TS 23.401 [11].

The targetMME or a target S4-SGSN extracts the source's NRI, RAC, LAC, MNC and MCC from the P-TMSI (or GUTI based on the procedure described in 3GPP TS 23.003 [4] sub-clause 2.8.2).

The FQDN based on NRI, RAC, LAC, MNC and MCC as defined in 3GPP TS 23.003 [4] sub-clause 19.4.2.6 is denoted in this specification as the NRI-RAI FQDN.

If the SGSN canonical node name employed by the operator is the NRI-RAI FQDN then see sub-clause 4.3.3.2 for provisioning.

If the 3GPP defined NRI-RAI FQDN is not employed by an operator as the SGSN's canonical node name then the operator shall provision NAPTR records under the NRI-RAI FQDN with at least "x-3gpp-sgsn:x-s3", "x-3gpp-sgsn:x-s4", "x-3gpp-sgsn:x-s16" (assuming the SGSN supports S3/S4/S16 GERAN/U-TRAN) and at least "x-3gpp-sgsn:x-gn", "x-3gpp-sgsn:x-gp" (assuming the SGSN supports legacy Gn/Gp).

The S-NAPTR procedure for finding the old SGSN services and interfaces from the P-TMSI is started with "Service Parameters" of

"x-3gpp-sgsn:x-gn", "x-3gpp-sgsn:x-gp", "x-3gpp-sgsn:x-s3", "x-3gpp-sgsn:x- s16"

as defined in 3GPP TS 23.003 [4] and setting the Application-Unique String to the NRI-RAI FQDN based on NRI, RAC, LAC, MNC and MCC as defined in 3GPP TS 23.003 [4] sub-clause 19.4.2.6:

NOTE 1: In the event a valid NRI is not available then the <NRI> value shall be excluded from the FQDN. The default SGSN in the SGSN pool shall be provisioned under that record (or the sole SGSN if there is no SGSN pool for that RAI).

NOTE 2: Service Parameters are logically limited to those supported by the target node that is performing the search for the source SGSN. So for example, a target Release-8 SGSN supporting only Gn/Gp would employ "x-3gpp-sgsn:x-gn" and "x-3gpp-sgsn:x-gp". An S4-SGSN only supporting S4/S3/S16 would employ "x-3gpp-sgsn:x-s3", "x-3gpp-sgsn:x-s16". A target MME would employ "x-3gpp-sgsn:x-s3" and may additionally include "x-3gpp-sgsn:x-gn" and "x-3gpp-sgsn:x-gp", to support the procedures in Annex D of 3GPP TS 23.401 [11].

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This list can be obtained one host name at a time, in a procedure similar to Annex C.2, or a complete ordered list of all nodes, in a procedure similar to Annex C.3. Such a complete list obtained from an S-NAPTR procedure is referred here as a candidate list.

NOTE 3: The services based on canonical node name can return services not starting with x-3gpp-sgsn since the SGSN node might, for example, have a co-located MME.

For a pre-Release-8 target node i.e. a UE moving from eUTRAN to pre-Release-8 UTRAN/GERAN the UE will provide a derived P-TMSI based on a GUTI (See Annex H of 3GPP TS 23.401 [11]). As a result the source MME or Release-8 SGSN looks like a pre-Release-8 SGSN to a pre-Release-8 target node. For pre-Release-8 compatibility operators would continue to provision A/AAAA records as described in Annex C.1 of 3GPP TS 23.003 [4] for the corresponding Gn/Gp interfaces regardless of whether the source SGSN is pre-Release-8 or not.

NOTE 4: Gn/Gp interfaces are provisioned redundantly for both ".gprs" and ".3gppnetwork.org" top level domains during the transition to Release-8 to allow a gradual forward migration to 3ggpnetwork.org while still supporting existing pre-Release-8 usage.

4.3.3.6 Services of a SGW from SGW canonical node name

An MME or S4-SGSN may need to find SGW interfaces on a SGW based solely on SGW's canonical node name. The most common use cases are:

- An MME finding an S11 interface from an SGW node name where the SGW node name was determined from an S5/S8 interface selection based on TAI (see sub-clause 5.2).
- An S4-SGSN finding an S4 interface from an SGW node name where the SGW node name was determined from an S5/S8 interface selection based on TAI (see sub-clause 5.2).
- Finding if a SGW node has PGW interfaces from an SGW node name (both SGW and PGW functions would be listed under one canonical node name for co-located PGW/SGW).

See sub-clause 4.3.3.2 for DNS provisioning of the canonical node name records.

For LTE initial attach cases, the S11 interface is initially unknown by an MME. The S5/S8 interface and the SGW hostname will be selected by procedures in subclause 5. The MME will obtain SGW S11 interfaces from the SGW canonical node name. The S-NAPTR procedure shall use "Service Parameter" of

"x-3gpp-sgw:x-11"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the canonical node name of the specific SGW node to find the available S11 interfaces

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of services and interfaces of that SGW (see Annex C.2 for a more detailed description of a candidate list).

For example, an operator might provision a SGW name at:

gw21.sgw.nodes.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

Similarly, for GERAN/UTRAN initial attach cases the S4-SGSN will need to obtain the SGW S4 interface after procedures in subclause 5 select the S5/S8 interface and SGW hostname. The only change from the MME case is the "Service Parameter" of

"x-3gpp-sgw:x-s4"

is employed.

In cases where a new PDN connection is added to an existing SGW the available SGW S5/S8 interfaces are commonly needed.

To resolve the allowed SGW PMIPv6 interfaces the S-NAPTR procedure shall be used with the "Service Parameters" of

"x-3gpp-sgw:x-s5-pmip", "x-3gpp-sgw:x-s8-pmip"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the canonical node name of the specific SGW node

Similarly for the S5/S8 GTP interfaces the S-NAPTR procedure shall use "Service Parameters" of

"x-3gpp-sgw:x-s5-gtp", "x-3gpp-pgw:x-s8-gtp"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the FQDN of the specific SGW node.

It is also possible to combine the "Service Parameters" in the S-NAPTR or to leave "Service Parameters" as logically unspecified initially in the S-NAPTR procedure in order to identify all interfaces for all 3GPP TS 29.303 supported protocols of the node.

NOTE 1: The services based on canonical node name can return services not starting with x-3gpp-sgw since the SGW node might, for example, have a co-located PGW.

5 Procedures for EPC Node Discovery and Selection

5.1 Procedures for Discovering and Selecting a PGW

5.1.1 Discovering a PGW for a 3GPP Access

5.1.1.1 General

The procedures here give a list of possible PGWs and their interfaces that serve a particular APN. This is very similar to the existing function that resolves the GGSN IP address based on an APN.

However, the Release-8 behaviour includes more functionality than pre-Release-8 systems since the PGW now can support more than one protocol and there is sometimes a desire to have the PGW and SGW collocated or topologically close to each other (with respect to the network topology), if possible. New DNS records are required to distinguish between different protocols and interfaces and assist in the more complicated selections.

The operator shall provision the authoritative DNS server responsible for the "apn.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org" domain with NAPTR records for the given APN-FQDN and corresponding PGWs under the APN-FQDN:

<APN-NI>.apn.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

See subclause 19.4.2.2 of 3GPP TS 23.003 [4].

The above format is used in DNS for use in DNS queries by S4-SGSN and MME to networks with DNS provisioned to Release-8. A Release-8 SGSN only supporting Gn/Gp may also optionally use this procedure.

The DNS records provisioned at that location are NAPTR records and include all S5/S8 and Gn/Gp interfaces for PGW, GGSN, and collocated PGW/GGSN that are intended to be used for that APN.

The pre-Release-8 format of

<APN-NI>.mnc<MNC>.mcc<MCC>.gprs

is still used in pre-Release-8 SGSN DNS queries and continues to be used as a fallback in Release-8 SGSN for discovering Gn/Gp interfaces in a pre-Release-8 network.

The DNS records provisioned at that location are A and/or AAAA records but only for the Gn/Gp interfaces of a standalone GGSN or collocated PGW/GGSN.

The APN-FQDN is derived from the APN where the APN is typically in the legacy format of <APN-NI>.mnc<MNC>.mcc<MCC>.gprs. as specified in sub-clause 9.1of 3GPP TS 23.003 [4].

NOTE 1: The APN-FQDN is used for DNS query purposes in Release-8. It does not imply a change in the use or format of the APN in other protocols, nodes or UE/MS. The APN-FQDN and the APN use independent formats but are related as below for DNS usage by the MME and S4-SGSN.

The APN received by the EPC node discovery function for 3GPP accesses, is always of the form of an APN-NI part and operator part. It is the output from Annex A of 3GPP TS 23.060 [18], which is exactly three labels with last label "gprs".

The APN is transformed to the APN-FQDN format of sub-clause 19.4.2.2.3 of 3GPP TS 23.003 [4] as follows.

a) If the APN has the last three labels matching the pattern "mnc<MNC>.mcc<MCC>.gprs", where <MNC> and <MNC> are each composed of 3 decimal digits, then the last 3 labels of the APN shall be replaced by "apn.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org" to form the APN-FQDN.

b) Else if the UE/MS is in the home network and the APN string has as last label "gprs", then last label "gprs" in the APN string shall be replaced by "apn.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org" to form the APN-FQDN where the MNC and MCC values are the values of the home network. This use case occurs if the HPLMN OI-1 in Annex A.2 **of** 3GPP TS 23.060 [18] does not conform to the pattern "mnc<MNC>.mcc<MCC>.gprs".

c) Otherwise the APN is invalid and cannot be used for a Release 8 APN usage.

In Annex A of 3GPP TS 23.060 [18] the SDL diagram refers to a "DNS interrogation" succeeding or failing which is the only DNS interaction. This is clarified as follows:

For the procedures defined in the present document the APN-FQDN shall be used in the S-NAPTR with a NAPTR query (see later subclauses for details). If the S-NAPTR procedure succeeds the "DNS interrogation" succeeds. If the S-NAPTR procedure fails to find a PGW or collocated PGW/GGSN then the MME (or an S4-SGSN that does not support Gn/Gp) then the "DNS interrogation" fails.

For the legacy procedures defined in Annex A of 3GPP TS 23.060 [18] the unmodified APN shall be used in the DNS A query and DNS AAAA query. If the A and AAAA records both fail then the MME would fail the DNS interrogation. If either query succeeds, the "DNS interrogation" succeeds.

5.1.1.2 Discovering a PGW or collocated PGW/GGSN for a 3GPP Access - S8/Gp roaming case existing PDN

This section covers the case where the SGW or S4-SGSN is in the visiting network, the SGW is already pre-selected by having at least one existing PDN connection and a UE attempts to create a new PDN connection for a different APN to be selected in the home network.

Operators shall provision NAPTR records for each APN-FQDN that allows roaming with at least "Service Parameters" of

"x-3gpp-pgw:x-s8-gtp", "x-3gpp-pgw:x-s8-pmip", "x-3gpp-ggsn:x-gp", "x-3gpp-pgw:x-gp"

for each such supported interface of that type.

The S-NAPTR procedure, employed by the MME or S4-SGSN, to discover all S8interfaces shall use "Service Parameters" of

"x-3gpp-pgw:x-s8-gtp", "x-3gpp-pgw:x-s8-pmip""

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and set the Application-Unique String to the APN FQDN as defined in subclause 19.4.2.2 of 3GPP TS 23.003 [4]. The "Service Parameter" of "x-3gpp-pgw:x-gp" shall be included if the MME or S4-SGSN wishes to potentially bias towards a collocated PGW/GGSN.

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of PGW or collocated PGW/GGSN for that APN (see Annex C.2 for an informative description of a candidate list and Annex B for the S-NAPTR procedure).

The above procedure shall be used by the MME or S4-SGSN to select the PGW or collocated PGW/GGSN.

NOTE 1: When an LTE capable terminal is in GERAN/UTRAN access, the S4-SGSN might wish to preferentially select a node with both Gp and S8 (i.e. a co-located PGW/GGSN) based on an operator policy. A preference for a co-located PGW/GGSN may also exist in an LTE access based on operator policy. The MME, or Release-8 SGSN, may find co-located PGW/GGSN nodes by searching the APN "candidate" list for interfaces with the same canonical node name in a Gp interface host name and S8 interface host name.

The PGW and SGW cannot be collocated in this case since the SGW and PGW are in different operator networks. Furthermore, topological matching by DNS host names shall not be done since the host names are under different operators' administrative control.

The Service Parameter of "x-3gpp-pgw:x-gp" denotes a collocated Release 8 GGSN function on a PGW. A PGW with a collocated Release 8 GGSN function may be preferred subject to operator policies. If that is the case the collocated PGW/GGSN nodes should be moved to the front of the candidate list but otherwise retaining the same relative order. The interfaces from the candidate list that are not S8 based shall be removed. The PGW S8 interfaces are tried in order from the candidate list.

NOTE 2: Contrary to the non-roaming case, in the roaming case the domain name of the SGW interface selected does not influence the PGW selection.

In the above procedure after the PGW has been contacted, the selected PGW node name, selected IP address, port (if non standard) and selected protocol type (GTPv2 vs. PMIP) shall be stored in the MME or S4-SGSN so it can be accessed on a PDN basis.

NOTE 3: In this release of 3GPP only standard ports are used.

3GPP TS 23.401 [11] currently indicates only one of PMIP or GTPv2 will be used based on roaming agreements so the above query would actually not require both *gtp* and *pmip*. The operator could use the order field in the NAPTR records to accomplish an optional fallback to the other protocol type.

Use cases where a SGW needs to be selected are covered in sub-clause 5.2. However, since Gn/Gp access bypasses SGW selection completely both for subsequent PDP context activations and initial attach we note that special case here.

If the UE is in GERAN or UTRAN access and the Release 8 SGSN supports Gp, but not S4, the procedure above is modified as follows. The "Service Parameters" shall be

"x-3gpp-pgw:x-gp" , "x-3gpp-ggsn:x-gp"

If an LTE capable mobile is in GERAN/UTRAN access a PGW with a collocated PGW/GGSN function may be preferred subject to operator policies. If that is the case the PGW/GGSN nodes should be moved to the front of the candidate list but otherwise retaining the same relative order. The rest of the procedure is the same as above.

If the APN record does not exist at the .3gppnetwork.org domain and the UE is in GERAN or UTRAN access and the Release 8 SGSN supports Gp then the pre Release-8 DNS procedures shall apply for the APN lookup by the Release 8 SGSN (i.e. APN lookup by A/AAAA records in the domain .gprs).

5.1.1.3 Discovering a PGW or collocated PGW/GGSN for a 3GPP Access - S5/Gn intra-operator existing PDN

Operators shall provision NAPTR records for each APN-FQDN for use within their network with at least "Service Parameters" of

"x-3gpp-pgw:x-s5-gtp", "x-3gpp-pgw:x-s5-pmip", "x-3gpp-ggsn:x-gn", "x-3gpp-pgw:x-gn"

for each such supported interface of that type.

Assuming the SGW is already pre-selected by having an existing PDN connection and a UE attempts to create a new PDN connection for a different APN in the user's home network, then the MME or S4-SGSN shall perform the following procedure:

The S-NAPTR procedure, employed by the MME or S4-SGSN to discover S5 interfaces shall use "Service Parameters" of

"x-3gpp-pgw:x-s5-gtp", "x-3gpp-pgw:x-s5-pmip"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and set the Application-Unique String to the APN FQDN as defined in subclause 19.4.2.2 of 3GPP TS 23.003 [4]. The "Service Parameter" of "x-3gpp-pgw:x-gn" shall be included if the MME or S4-SGSN wishes to potentially bias towards a collocated PGW/GGSN.

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of PGW or collocated PGW/GGSN for that APN (see Annex C.2 for an informative description of a candidate list and Annex B for the S-NAPTR procedure).

Collocation and topological ordering between the PGW and SGW applies in this case.

If the existing SGW hostname has "topoff" then the candidate list of PGW shall be used in the order given to try to contact a PGW, after moving any colocated SGW/PGW to the front of the candidate list while maintaining relative order within that set.

The Service Parameter of "x-3gpp-pgw:x-gn" denotes a collocated Release 8 GGSN function on a PGW. A PGW with a collocated Release 8 GGSN function may be preferred subject to operator policies. If that is the case the collocated PGW/GGSN nodes should be moved to the front of the candidate list but otherwise retaining the same relative order. The interfaces from the candidate list that are not S5 based shall be removed. The PGW S5 interfaces are tried in order from the candidate list..

If the existing SGW hostname has "topon" the two candidate lists shall be used in the procedure in Annex C.4 with the PGW as "A" and the SGW as "B". Annex C.4 results in a list of PGW to try in order.

Once a PGW is successfully contacted the selected PGW host name, PGW IP address used, port (if non-standard) and selected protocol type (GTP vs PMIP) shall be stored in the MME or S4-SGSN so it can be accessed on a PDN basis.

NOTE 1: In this release of 3GPP only standard ports are used.

Use cases where a SGW needs to be selected are covered in sub-clause 5.2 and sub-clause 5.3. However, since Gn/Gp access bypasses SGW selection completely both for subsequent PDP context activations and initial attach we note that special case here.

If the UE is in GERAN or UTRAN access and the Release 8 SGSN supports Gn, but not S4, the procedure above is modified as follows. The "Service Parameters" shall be

"x-3gpp-pgw:x-gn" , "x-3gpp-ggsn:x-gn"

If an LTE capable mobile is in GERAN/UTRAN access a PGW with a collocated PGW/GGSN function may be preferred subject to operator policies. A preference for a co-located PGW/GGSN may also exist in an LTE access based on operator policy. The MME, or Release-8 SGSN, may find co-located PGW/GGSN nodes by searching the APN "candidate" list for interfaces with the same canonical node name in a Gn interface host name and S5 interface host name. If that is the case the PGW/GGSN nodes should be moved to the front of the candidate list but otherwise retaining the same relative order. The rest of the procedure is the same as above.

If the APN record does not exist at the .3gppnetwork.org domain and the UE is in GERAN or UTRAN access and the Release 8 SGSN supports Gn then the pre Release-8 DNS procedures shall apply for the APN lookup by the Release 8 SGSN (i.e. APN lookup by A/AAAA records in the domain .gprs)

5.1.1.4 Discovering a PGW, collocated PGW/GGSN or GGSN for a 3GPP Access - S5/Gn intra-operator initial attach

During the initial attach and PDN connection creation using a 3GPP access both a PGW and an SGW need to be selected by an MME and will also be used by a S4-SGSN in PDP context creation. The discovery and selection procedures for cases employing a SGW are the same as for the PGW and the SGW discovery and selection procedure described in subclause 5.3.

The discovery and selection procedures for a Release-8 SGSN selecting a Gn interface for PDP context creation are in subclause 5.1.1.3.

5.1.2 Discovering a PGW for a non-3GPP Access with Network Based Mobility Management

5.1.2.1 Discovering a PGW for a non-3GPP Access – S2a/S2b initial attach for roaming and non-roaming

The MAG functionality within the trusted non-3GPP IP access or the e-PDG shall use the S-NAPTR procedure with "Service Parameters" of

"x-3gpp-pgw:x-s2a-pmip", "x-3gpp-pgw:x-s2b-pmip", "x-3gpp-pgw:x-s2a-mipv4"

and the APN-FQDN as the Application-Unique String.

<APN-NI>.apn.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of PGW for that APN (see Annex B for S-NAPTR procedure and see Annex C.2 for an informative description of a candidate list).

There is no requirement for selection for a collocated PGW/SGW in this procedure. In the above procedure the selected PGW node name, port and selected type (PMIPv6, MIPv4) shall be stored in the MAG so it can be accessed on a PDN basis.

5.1.2.2 Discovering a PGW for a non-3GPP Access – S2a/S2b initial attach and chained S2a/S2b with GTP or PMIPv6 based S8

The MAG functionality within the trusted non-3GPP IP access or the e-PDG shall use the S-NAPTR procedure with the "Service Parameters" of

"x-3gpp-pgw:x-s2a-pmip", "x-3gpp-pgw:x-s2b-pmip"

and the APN-FQDN as the Application-Unique String.

<APN-NI>.apn.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of PGW for that APN (see Annex B for S-NAPTR procedure and see Annex C.2 for an informative description of a candidate list).

The MAG selects a PGW based on the protocol type (GTP v/s PMIPv6) supported over the S5/ S8 interface based on information received over STa and SWm interfaces.

The PGW and SGW cannot be collocated in this case since the SGW and PGW are in different operator networks.

The DNS records in the order returned are then used to contact the PGW node.

5.1.3 Discovering a PGW for a non-3GPP Access with DSMIPv6

5.1.3.1 Discovering a PGW for a non-3GPP Access – S2c initial attach

This section covers the case where the IP address of the Home Agent (HA) functionality of a particular PGW needs to be discovered from the FQDN of the PGW. This query may be sent from a trusted access gateway or from an ePDG. The trusted access gateway or ePDG shall use the S-NAPTR procedure with "Service Parameters" of

"x-3gpp-pgw:x-s2c-dsmip"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the FQDN of the specific PGW.

5.2 Procedures for Discovering and Selecting a SGW

5.2.1 General

These procedures are employed when an SGW needs to be selected by an EPC core node and a PGW has already been selected. In particular for the tracking area update procedure with SGW change.

The SGW is selected based on the target cell where the UE has moved into. The MME has the new target eNodeB cell ID (eCID) and TAI available . The MME shall construct the TAI FQDN as defined in subclause 19.4.2.3 of 3GPP TS 23.003 [4].

Operators shall provision, for each TAI value in their network, NAPTR records under the TAI FQDN corresponding to each valid SGW interfaces from the following "Service Parameters"

"x-3gpp-sgw:x-s8-gtp", "x-3gpp-sgw:x-s8-pmip", "x-3gpp-sgw:x-s5-gtp", "x-3gpp-sgw:x-s5-pmip"

where additional "Service Parameters" may be included optionally.

For each RAI value that is served by a S4-SGSN the same records would be provisioned under the RAI FQDN (see sub-clause 5.5.2 for the RAI FQDN).

The S-NAPTR procedure employed by an MME for finding a candidate set of SGW nodes shall use the TAI FQDN as the Application-Unique String.

For the purposes of this document the NAPTR record-set at that location will be called the TAI NAPTR record-set.

The MME selects the S11 interface of the SGW from the SGW's canonical node record (see subclause 4.3.3) if it is not obtained from the TAI records.

NOTE: If an operator does not use the "a" and "s" flags in the TAI NAPTR records (i.e. they use the "" flag) and they are using SGW service areas it is strongly recommended that the TAI NAPTR records point directly to NAPTR records representing the SGW service areas. This is to facilitate possible future use in the SGW Load Re-balancing procedure.

For the case of an S4-SGSN making the SGW selection the RAI FQDN (see sub-clause 5.5.2) is used instead of the TAI FQDN to select the SGW but is otherwise the same as the MME handling TAU.. The S4-SGSN selects the S4 interface of the SGW from the SGW's canonical node record (see subclause 4.3.3) if it is not obtained from the RAI records.

S-GW selection when SGW that acts as a local anchor for non-3GPP access in the case of S8-S2a/b chained roaming is outside the scope of this specification.

5.2.2 SGW Selection during TAU or RAU with SGW change - 3GPP roaming case

For the roaming case the type of protocol (PMIP vs. GTP) is chosen based on a roaming agreement according to 3GPP TS 23.401 [11]. The MME shall therefore use the S-NAPTR procedure with "Service Parameters" of

"x-3gpp-sgw:x-s8-gtp" or "x-3gpp-sgw:x-s8-pmip"

(based on whetherGTP v2 or PMIPv6 is used for the current PDN connection)

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and possibly restricted based ona roaming agreement to use only GTPv2 or PMIPv6and set the Application-Unique String to the TAI FQDN as defined in subclause 19.4.2.3 of 3GPP TS 23.003 [4].

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of SGW for that TAI (see Annex B for S-NAPTR procedure and see Annex C.2 for an informative description of a candidate list).

If the first choice protocol (PMIP or GTPv2) fails the second choice MAY be tried subject to the operators' roaming agreements.

Neither collocation of PGW and SGW nor topological ordering rules apply in this case.

The present sub-clause to this point implictly assumes only one PDN connection is currently employed by a UE which may not be the case with multiple PDN connections for the same UE. All PDN connections after the TAU must be on the same SGW since there is only one SGW at a time for a UE.

If an existing PDN for the UE is PMIPv6 S8 and some existing PDN for the UE is GTPv2 S8 then a SGW supporting both protocols is required. If that occurs and there are no such SGW then the PDN connections with the least important retention (from the ARP value) would have to be dropped until a list of viable SGW of one protocol meeting the retention policies is found. If there are non viable SGW they are removed from the original SGW candidate list. After this point the SGW candidate list order is used in the same way as it was in the case of only one PDN connection.

Once an SGW is successfully contacted the selected SGW host name, selected SGW IP address, selected port (if non-standard) and selected protocol type (GTPv2 vs. PMIP - which is unchanged here) shall be stored in the MME so it can be accessed on a PDN basis.

NOTE 1: In this release of 3GPP only standard ports are used.

For the case of an S4-SGSN making the SGW selection the RAI FQDN (see sub-clause 5.5.2) is used instead of the TAI FQDN to select the SGW but is otherwise the same as the MME handling TAU. The S4-SGSN selects the S4 interface of the SGW from the SGW's canonical node record (see subclause 4.3.3) if it is not obtained from the RAI records.

NOTE 2: A SGW will need to be selected at a handover attach from another access type. The SGW selection method is the same as presented here since there are existing PDN connections (typically PMIPv6 for non-3GPP accesses).

5.2.3 SGW Selection during TAU or RAU with SGW change - non-roaming case

This differs from the 3GPP roaming case in 5.2.2 primarily in that the PGW and the SGW are in the same network. Hence, there is a need to be able of selecting a SGW collocated with the PGW or a topologically close SGW. The current PGW's node name should previously have been stored in the MME or S4-SGSN when the default bearer was established or transported from the old MME to the new MME, and is therefore available for comparison.

For the non-roaming case the S-NAPTR procedure shall be initiated with "Service Parameters" of

"x-3gpp-sgw:x-s5-gtp" and/or "x-3gpp-sgw:x-s5-pmip"

(based on whether GTPv2 or PMIPv6 is used for the current PDN connection)

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and set the Application-Unique String to the TAI FQDN as defined in subclause 19.4.2.3 of 3GPP TS 23.003 [4].

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of SGW for that TAI (see Annex B for S-NAPTR procedure and see Annex C.2 for a more detailed informative description of a candidate list).

Collocation of PGW and SGW and topological ordering rules both apply in this case.If the existing PGW hostname for the PDN has "topoff" then the "candidate" list of SGW would be used in the order given to try to contact a SGW after moving the PGW with the same SGW node name to the front of the list keeping relative order..

If the existing PGW hostname has "topon" the two candidate lists shall be used in the procedure in Annex C.4 with the SGW as "A" and the PGW as "B". Annex C.4 results in a list of SGW to try in order.

The present sub-clause to this point implictly assumes only one PDN connection is currently employed by a UE which may not be the case with multiple PDN connections for the same UE.

If an existing PDN for the UE is PMIPv6 S5 and some existing PDN for the UE is GTPv2 S5 then a SGW supporting both protocols is required. If that occurs and there are no such SGW then the PDN connections with the least important retention (from the ARP value) would have to be dropped until a list of viable SGW of one protocol meeting the retention policies is found. If there are non viable SGW they are removed from the original SGW candidate list. After this point the SGW candidate list order is used in the same way as it was in the case of only one PDN connection.

After this point it is a matter of operator policy or vendor implementation which PDN or PDNs (and hence their corresponding PGW canonical node names) are used for selecting the corresponding best SGW interface.

NOTE 1: One possible option would be as follows. First try to maximize the number of PDN connections that are on PGW colocated on the same viable SGW. If there are no collocated choices the PGW giving the closest topological match to any viable SGW are used. If they are all equal then SGW ordering from the SGW candidate list is used. This approach would be very similar, but not identical, to passing the list of PGW being used as list "B" and the SGW candidate list with only viable SGW as list "A' in the procedure in Annex C.4

Once an SGW is successfully contacted the selected SGW host name, SGW IP address used, port (if non-standard) and selected type (GTP vs PMIP) shall be stored in the MME so it can be accessed on a PDN basis.

NOTE 2: In this release of 3GPP only standard ports are used.

For the case of an S4-SGSN making the SGW selection the RAI FQDN (see sub-clause 5.5.2) is used instead of the TAI FQDN to select the SGW but is otherwise the same as the MME handling TAU. The S4-SGSN selects the S4 interface of the SGW from the SGW's canonical node record (see subclause 4.3.3) if it is not obtained from the RAI records.

NOTE 3: A SGW will need to be selected at a handover attach from another access type. The SGW selection method is the same as presented here since there are existing PDN connections (typically PMIPv6 for non-3GPP accesses).

5.2.4 SGW Selection during non-3GPP handover to 3GPP access

The SGW selection is similar with other cases. For the non-roaming case, there is a need of selecting a SGW collocated with the PGW or a topologically close SGW. The current PGW's node name should previously have been stored in the HSS and would be sent to the MME during the access authentication procedure and is therefore available for comparison.

The S-NAPTR procedure for SGW Selection is same as section 5.2.2 for roaming case, and is same as section 5.2.3 for non-roaming case.

5.3 Procedures for Discovering and Selecting a PGW and SGW

This scenario applies to the UE initial attach case, where the MME or S4-SGSN has not yet assigned a PGW or a SGW to the UE. During the attach procedures, the MME shall select the SGW and the PGW as described below. During the UTRAN/GERAN attach procedure, the S4-SGSN shall select the SGW and the PGW as described below.

The selected SGW shall serve the UE's cell. During the attach procedure the MME receives the eNodeB cell ID, which has the TAI value.

The S-NAPTR procedure to obtain a list of "candidate" SGW shall be used with "Service Parameters" of

"x-3gpp-sgw:x-s5-gtp", "x-3gpp-sgw:x-s5-pmip"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and set the Application-Unique String set to the TAI FQDN as defined in subclause 19.4.2.3 of 3GPP TS 23.003 [4].

The S-NAPTR procedure logically outputs a list of host names each coupled with a service, a protocol, a port, and a list of IPv4 and IPv6 addresses. This is the "candidate" list of SGWs for a specific TAI (see Annex B for S-NAPTR procedure and see Annex C.2 for an informativedescription of the candidate list).

The list of "candidate" PGW is obtained as follows:

The S-NAPTR procedure to get the list of "candidate" PGW shall use "Service Parameters" of

"x-3gpp-pgw:x-s5-gtp", "x-3gpp-pgw:x-s5-pmip", "x-3gpp-pgw:x-gn"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4], and the Application-Unique String set to the APN FQDN as defined in subclause 19.4.2.2 of 3GPP TS 23.003 [4].

The S-NAPTR procedure logically outputs a list of host names each coupled with a service, a protocol, a port, and a list of IPv4 and IPv6 addresses. This is the "candidate" list of PGWs for a specific APN (see Annex C.2 for a detailed description of a candidate list).

The two candidate lists shall be used in the procedure described in Annex C.4 with the SGW as an "A" node type and the PGW as a "B" node type in the procedure.

The procedure described in Annex C.4 results in a selection of a SGW and a PGW along with the protocol, the IP address and the port. In the case of a failure to contact the SGW or the PGW, the required gateway reselection procedures are described in Annex C.4.

The MME selects the S11 interface of the SGW from the SGW's canonical node record (see subclause 4.3.3) if it is not obtained from the TAI records.

NOTE: The MME (S4-SGSN) send a GTPv2 Create Session Request to the SGW respectively over S11 or S4 with the IPv4/IPv6 address of the PGW. After the SGW has been successfully contacted over S11 or S4, the SGW can try to contact the PGW over S5/S8.

Once the SGW has successfully been contacted, the selected SGW host name, the used SGW IP address, the port number and the selected protocol type shall be stored in the MME or S4-SGSN per PDN.

Once the PGW has successfully been contacted, the selected PGW host name, the used PGW IP address, the port number and the selected protocol type shall be stored in the MME or S4-SGSN per PDN.

For the case of an S4-SGSN making the selection the RAI FQDN is used instead of the TAI FQDN to select the SGW but is otherwise the same as an MME doing the selection. The S4-SGSN selects the S4 interface of the SGW from the SGW's canonical node record (see subclause 4.3.3) if it is not obtained from the RAI records.

5.4 Procedures for Discovering and Selecting an MME

These procedures can be used as part of the 'Inter eNodeB handover with MME relocation' procedure to select an MME.

The MME is selected based on the target cell where the UE has just moved into. The MME has the new target eNodeB cell ID (eCID) and TAI available . The TAI includes the TAC, the MNC and the MCC.

The S-NAPTR procedure for finding a candidate set of MME nodes for a target TAI is always started at the TAI FQDN as defined in subclause 19.4.2.3 of 3GPP TS 23.003 [4]. The S-NAPTR procedure performed at the source MME for finding a candidate set of target MME nodes is started with at least the "Service Parameters" of

"x-3gpp-mme:x-s10"

as defined in subclause 19.4.3 of 3GPP TS 23.003 [4].

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of MME for that TAI (see Annex C.2 for a more detailed description of a candidate list).

Topological closeness shall not be used for the MME selection so the MME candidate list would be used in the order given to try to contact a MME.

NOTE 1: The source MME may wish to preferentially select a target MME that is co-located with a SGSN function based on an operator policy. The S-NAPTR procedure using the canonical node name (sub-clause 4.3.3) may be used to find any co-located SGSN interfaces in the candidate list. Optionally an operator may provision NAPTR records with "x-3gpp-sgsn:x-gn" and/or "x-3gpp-sgsn:x-s4" at the TAI FQDN to allow an MME to directly get a candidate list for the co-located SGSN and hence for the source MME to match node names from the hostnames in the two candidate lists.

NOTE 2: If an operator does not use the "a" and "s" flags in the TAI NAPTR records (i.e. they use the "" flag) it is strongly recommended that the TAI NAPTR records point directly to NAPTR records at mmegi<MMEGI>.mme.epc.mnc<mnc>. mcc<mcc>.3gppnetwork.org (i.e. an MME pool)

For an S4-SGSN making the selection the "Service Parameter" of "x-3gpp-mme:x-s3" shall be employed instead of "x-3gpp-mme:x-s10".

For the case when a UE is moving from a pre-Release-8 UTRAN/GERAN to an MME the pre-Release-8 source node will not be able to use the .3ggpnetwork.org based records. As a result the MME being selected looks like a pre-Release-8 SGSN to a pre-Release-8 node performing a selection. For pre-Release 8 compatibility, operators would continue to provision A/AAAA records as per Annex C.1 of 3GPP TS 23.003 [4] for the corresponding MME Gn/Gp interfaces under the RNC-ID corresponding to the mapped targed ID value (see subclause D.3.4 of 3GPP TS 23.401 [11] and subclause 5.5.3 of the present document).

5.5 Procedures for Discovering and Selecting an SGSN

5.5.1 General

These procedures are employed when a target SGSN that serves the target cell or RNC needs to be initially selected by an EPC core node or Release-8 SGSN. It explicitly does not cover selection of an SGSN by a RAN node.

EPC core nodes, i.e. the MME and the S4-SGSN employ the procedures below during the SRNS relocation procedure. A Release-8 SGSN supporting only Gn/Gp may also optionally use this procedure.

The SGSN is selected based on information in the Target ID (see 3GPP TS 23.003 [4] and 3GPP TS 25.413 [12]). For PS GERAN the Target ID has the global cell ID including PLMN and for U-TRAN the Target ID has RAC, RNC-ID and PLMN.

5.5.2 SGSN initial target selection based on RAI

In both U-TRAN and GERAN cases the target RAC, LAC, MNC, and MCC are available from the information in the Target ID. This selection is done by an MME or Release-8 SGSN during SRNS relocation procedures.

The S-NAPTR procedure for an MME finding a candidate set of SGSN services and interfaces serving the target Routing Area is started with "Service Parameters" of

"x-3gpp-sgsn:x-gn", "x-3gpp-sgsn:x-gp", "x-3gpp-sgsn:x-s3"

as defined in 3GPP TS 23.003 [4] and setting the Application-Unique String to the RAI FQDN based on RAC, LAC, MNC, MCC as defined in 3GPP TS 23.003 [4] sub-clause 19.4.2.5:

rac<RAC>.lac<LAC>.rac.epc.mnc<MNC>. mcc<MCC>.3gppnetwork.org

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of SGSN for that RAI (see Annex B for S-NAPTR procedure and see Annex C.2 for a more informative description of a candidate list).

NOTE 1: For an MME making the selection there is an existing PDN connection (and hence an existing SGW) so the S3 interfaces would logically be preferred over the Gn/Gp interfaces to keep the existing PDN connection. However, if there is a colocated PGW/GGSN being employed with the 3GPP TS 23.401 [11] Annex D procedures then the Gn/Gp interfaces on the target SGSN can also be valid. Operators should prioritize the DNS records according to their preference.

For an S4-SGSN making the selection the "Service Parameter" of "x-3gpp-sgsn:x-s16" shall be employed instead of "x-3gpp-sgsn:x-s3".

A Release 8 SGSN supporting only Gn/Gp may also optionally use this procedure with "Service Parameters" of "x-3gpp-sgsn:x-gn" and "x-3gpp-sgsn:x-gp".

Operators shall provision, for each RAI value in their network, NAPTR records under the RAI FQDN corresponding to each valid SGSN interfaces from at least the following "Service Parameters"

"x-3gpp-sgsn:x-gn", "x-3gpp-sgsn:x-gp", "x-3gpp-sgsn:x-s3","x-3gpp-sgsn:x-s16",

where NAPTR records for additional "Service Parameters" may be included optionally.

NOTE 2: The NAPTR record at the RAI FQDN can be provisioned to correspond to only the default SGSN node(s) in the SGSN pool(s) serving that RAI. The default SGSN after the DNS procedure exits would be contacted with GTP, the default SGSN then selects the actual SGSN within that SGSN pool. This results in all relocation requests being handled by the default SGSN nodes. If the operator's goal is to avoid load on the default SGSN nodes then the records provisioned at the RAI FQDN should instead include the entire set of SGSN in all SGSN pools that service that RAI. The S-NAPTR procedure would then return each SGSN in the SGSN pool based on the provisioned DNS weights and priority.

NOTE 3: The SGSN(s) closest to the geographical region covered by the RAI can be biased by provisioning the DNS records with higher priority or weights.

3GPP does not require that collocation and "topon" naming is applicable in SGSN selection.

NOTE 4: Service parameters are limited to those supported by the node doing the search.

NOTE 5: SGW records will also be provisioned under the RAI FQDN see subclause 5.2

For the case when a UE is moving from a pre-Release-8 UTRAN/GERAN to a Release-8 target SGSN the pre-Release-8 source node will not be able to use the .3ggpnetwork.org based records. As a result the Release 8 SGSN (or MME) being selected looks like a pre-Release 8 SGSN to a pre-Release-8 node performing a selection. For pre-Release 8 compatibility operators would continue to provision A/AAAA records as per Annex C.1 of 3GPP TS 23.003 [4] for the corresponding Gn/Gp interfaces regardless of whether the SGSN is pre-Release-8 or not. Vendors shall support pre-Release 8 DNS procedures on Release 8 SGSN for DNS queries to pre-Release 8 networks or nodes.

5.5.3 SGSN initial target selection based on RNC-ID (UTRAN target)

NOTE 1: The finer granularity this procedure allows only applies to UTRAN and only when different RNC-IDs have the same RAI values.

This procedure is used only for a UTRAN target in the SRNS procedure.

In UTRAN case the target RNC-ID, MNC, and MCC are available from the information in the Target ID.

The S-NAPTR procedure for an MME finding a candidate set of SGSN services and interfaces serving the target RNC is started with "Service Parameters" of

"x-3gpp-sgsn:x-gn", "x-3gpp-sgsn:x-gp", "x-3gpp-sgsn:x-s3"

as defined in 3GPP TS 23.003 [4] and setting the Application-Unique String to the RNC-ID FQDN based on RNC-ID,MNC,MCC as defined in 3GPP TS 23.003 [4] sub-clause 19.4.2.7.

The S-NAPTR procedure logically outputs a list of host names each with a service, protocol, port and a list of IPv4 and IPv6 addresses. This is a "candidate" list of SGSN for that RNC-ID (see Annex B for S-NAPTR procedure and see Annex C.2 for a more informative description of a candidate list).

NOTE 2: For an MME making the selection there is an existing PDN connection (and hence an existing SGW) so the S3 interfaces would logically be preferred over the Gn/Gp interfaces to keep the existing PDN connection. However, if there is a colocated PGW/GGSN being employed with the 3GPP TS 23.401 [11] Annex D procedures then the Gn/Gp interfaces on the target SGSN can also be valid. Operators should prioritize the DNS records accordingly.

For an S4-SGSN making the selection the "Service Parameter" of "x-3gpp-sgsn:x-s16" shall be employed instead of "x-3gpp-sgsn:x-s3".

A Release 8 SGSN supporting only Gn/Gp may also optionally use this procedure with "Service Parameters" of "x-3gpp-sgsn:x-gn" and "x-3gpp-sgsn:x-gp".If there are no NAPTR records at that RNC-ID then the RAI based lookup in subclause 5.5.2 shall be used as a fallback.

Operators using this feature shall provision, for each RNC-ID value in their network, NAPTR records under the RNC-ID FQDN corresponding to each valid SGSN interfaces from the following "Service Parameters"

"x-3gpp-sgsn:x-gn", "x-3gpp-sgsn:x-gp", "x-3gpp-sgsn:x-s3", "x-3gpp-sgsn:x-s16",

where NAPTR records for additional "Service Parameters" may be included optionally.

Operators not using this feature shall not provision the RNC-ID records.

NOTE 3: The NAPTR record at the RNC-ID FQDN can be provisioned to correspond to only the default SGSN node(s) in the SGSN pool(s) serving that RNC. The default SGSN after the DNS procedure exits would be contacted with GTP, the default SGSN then selects the actual SGSN within that SGSN pool. This results in all relocation requests being handled by the default SGSN nodes. If the operator's goal is to avoid load on the default SGSN nodes then the records provisioned at the RNC-ID FQDN should instead include the entire set of SGSN in all SGSN pools that service that RNC. The S-NAPTR procedure would then return each SGSN in the SGSN pool based on the provisioned DNS weights and priority.

NOTE 4: The SGSN(s) closest to the geographical region serving the RNC can be biased by provisioning the DNS records with higher priority or weights.

3GPP does not require that collocation and "topon" naming is applicable in SGSN selection.

NOTE 5: Service parameters are limited to those supported by the node doing the search.

For the case when a UE is moving from a pre-Release-8 UTRAN/GERAN to a Release-8 target SGSN the pre-Release-8 source node will not be able to use the .3ggpnetwork.org based records. As a result the target Release-8 SGSN (or MME) looks like a pre-Release-8 SGSN to a pre-Release-8 source node. For pre-Release 8 compatibility operators would continue to provision A/AAAA records as per Annex C.3 of 3GPP TS 23.003 [4] for the corresponding Gn/Gp interfaces regardless of whether the target SGSN is pre-Release-8 or not. Vendors shall support pre-Release 8 DNS procedures on Release 8 SGSN for DNS queries to pre-Release 8 networks or nodes.

Annex A (Informative): Examples

A.1 Introduction

This annex includes examples of the DNS provisioning needed for EPC node discovery and selection procedures. Examples are not exhaustive either in scope or in content but are intended to be sufficient to illustrate the general techniques and some of the more important use cases.

A.2 Preconditions

The following are general comments applying to all examples provided here.

a) DNS master file format follows ITEF RFC 1035 [3] chapter 5. The reader is reminded that in that format the character ":" is used for comments, "$" is used for directives, and enclosing parenthesis () combine lines.

b) Examples are built from smaller files and merged with the $INCLUDE directive. This is to allow each file to be documented in a seperate subclause and to allow reuse of files in some of the examples.

c) Only one operator's network is used in all examples using MCC=311 (United States) and MNC=990 (currently an unassigned MNC value). All DNS records presented here are within the zone cut epc.mnc990.mcc311.3gppnetwork.org.

d) Since the fully qualified domain names here are relatively long to avoid line wrapping and to save space in the examples we use the standard $ORIGIN directive
$ORIGIN epc.mnc990.mcc311.3gppnetwork.org.
Hence any domain names in the file that are not fully qualified (i.e. does not end with a period) would have epc.mnc990.mcc311.3gppnetwork.org. appended to them to form the fully qualified value. In the descriptive text in this Annex we use the same convention.

e) When a printout is provided, it is from an actual printout of the "dig" command (from the BIND distribution) but with all values of the string .epc.mnc990.mcc311.3gppnetwork.org. replaced with .$ORIGIN in order to keep the output on one line when possible. Note the "+tcp" option is used in the "dig" command to avoid truncating large DNS responses in the examples.

f) Files provided here have been tested with a recent version of BIND and are presented here unmodified. In the examples the BIND server was set as authoritative only with default settings to make the configuration self contained. In a real network a recursive server is likely to be used as an intermediary and/or multiple authoritative servers.

g) The time to live (TTL) in the examples are all set to 1 hour = 3600 seconds by using the $TTL directive from IETF RFC 2038 [15].

h) IPv4 addresses in the examples are from the IPv4 example documentation address space of 192.0.2.0/24 of IETF RFC 3330 [16] and IPv6 addresses are from the IPv6 documentation address space of IETF RFC 3849 [17] . These are by no means representative and are used solely for example purposes.

i) Since A/AAAA records are well known we only provide A/AAAA records for the simpler examples and do not document every A and AAAA lookup in detail.

A.3 Colocated Simple LTE Example

A.3.1 Network description

This is a fairly complete example of an operator network with the following properties.

1) Network only has LTE access with only interworking with LTE access. There is no support for Gn/Gp and no support for S3/S4.

2) Network does not support PMIPv6 for S5 but does allow very limited roaming use of S8 PMIPv6. This is a policy that is extremely unlikely to be used by a real operator and is used only for illustrative purposes.

3) There are two combined PGW/SGW nodes and two standalone MME nodes.

4) Any load balancing between nodes of the same functional type is on a strict bias. Hence, SRV records are not needed.

5) Operator desires the flattest DNS structure possible due to desire to minimize the number of DNS queries. The use of NATPR with flag "" is therefore not used by operator policy. This is feasable since the network is small.

6) Due to 4) and 5) only NAPTR records with flag "a" are used by the operator in this example.

7) One MME pool area.

8) One SGW service area.

9) Operator desires that S1-U link cost is minimized (i.e. SGW closest to eNodeB is used)

10) Operator does not care about distance or transport costs between SGW and PGW beyond the collocation for S5. Hence it does not use "topon" naming.

11) Small number of APNs

Other than item 2) this set of conditions might fit a green field operator or government agency deploying LTE for use on laptop computers or high end hand held devices for internet access. This represents one of the simpler networks possible while still being non-trivial.

Very large networks would tend to use MME pool areas and SGW service areas and NAPTR empty flag records. How complex the structure in DNS is employed between the initial NAPTR entry point and the final A/AAAA record set(s) is solely the operator's choice so long as the S-NAPTR requirements are met.

Additional comments on this example are in each section as needed. Also commented out DNS records are often included in the sample files to illustrate how additional interface types would be added.

A.4.2 Master file for "Collocated Simple LTE Example"

The master file for this example will be CS_EX.txt and has following content.

```
$ORIGIN epc.mnc990.mcc311.3gppnetwork.org.
;
$TTL    3600 ;  1 hour  - this directive is defined in IETF RFC 2308 not RFC 1035
;
$INCLUDE SOA_DB.txt
;
$INCLUDE CS_MME_DB.txt
;
$INCLUDE CS_SGW_PGW_NODE_DB.txt
;
$INCLUDE CS_TAI_DB.txt
;
$INCLUDE CS_APN_DB.txt
;
; End of file
```

The included files are detailed in following sections.

A.4.3 SOA and NS records

The SOA and NS records need to exist for any DNS zone. Our example will use the following file with name SOA_DB.txt whose content is:

```
;  SOA and NS glue records
;  Note @, ns1, ns2 are relative to $ORIGIN
;
```

```
@  1D  IN      SOA @ administrator.example.com. (
                     2008122401 ; serial
                     1H ; refresh
                     15 ; retry
                     1w ; expire
                     1h ; minimum - note this parameter
                        ; is used for negative caching
                        ; see RFC 2308 chapter 4
                    )
       IN  NS     ns1
       IN  NS     ns2
;
; Glue records for the two DNS servers
ns1   IN  A      192.0.2.247
ns2   IN  A      192.0.2.248

; End of file
```

A.4.4 MME file for "Collocated Simple LTE Example"

The MME S10 record lookup by GUTI is probably the simplest example of the NAPTR records and is the simplest introduction to the S-NAPTR procedure as well.

A NAPTR record is required to be provisioned at mmec<MMEC>.mmegi<MMEGI>.mme (see subclause 4.3.3.4 of this TS). This is used for GUTI based lookup of the old or source MME. Specifically, the target MME takes the UE's old GUTI value and derives the old MME's DNS name. The goal of the target MME is usually to be able to do a context transfer of the old MME's data (or similar action). MME to MME communication is done over S10. Hence, S10 is critical for the GUTI record in an LTE only network.

The mmec<MMEC>.mmegi<MMEGI>.mme is also the natural place for the MME canonical node record set. For this example we place the MME canonical node records and the required record for GUTI based lookup at the same location both for simplicity and better DNS caching.

For this example we have two MME with MMEC codes of 1 and 2 respectively and both in the same MME pool (MMEGI = 8000 hexadecimal).

The file containing the MME records for this example will be CS_MME_DB.txt and has following NAPTR record content.

```
;
; This is the GUTI related record and the MME node record
; The operator has decided to use the 3GPP name as the canononical node name of the MME
; rather than having two records (the 3GPP one and an operator defined value)
;
mmec01.mmegi8001.mme                        (
;     IN NAPTR order pref. flag service                    regexp  replacement
      IN NAPTR 100   999   "a"   "x-3gpp-mme:x-s10"           ""  topoff.eth1.mmec01.mmegi8001.mme )
      IN NAPTR 200   999   "a"   "x-3gpp-mme:x-s11"           ""  topoff.eth3.mmec01.mmegi8001.mme
;     IN NAPTR 300   999   "a"   "x-3gpp-mme:x-s3"            ""  topoff.eth5.mmec01.mmegi8001.mme
;     IN NAPTR 400   999   "a"   "x-3gpp-mme:x-gn"            ""  topoff.eth6.mmec01.mmegi8001.mme
;     IN NAPTR 500   999   "a"   "x-3gpp-mme:x-s1-mme"        ""  topoff.eth7.mmec01.mmegi8001.mme

;  This particular operator only supports LTE access in their accesses etc.
;  So the S3 record is commented out above. If the operator wants the MME to be used with S3/S4
;  SGSN nodes then the record would have to be included.
;  The Gn/Gp interface is commented out for same reason. This operator does not support it.
;  If the operator wants the MME to be used for Gn/Gp SGSN interworking
;  then the record would have to be included.
;
;   Reminder: Canonical node name records must be complete.
;   However, "x-3gpp-mme:x-s1-mme" is an exception.
;   While S1-MME interface must be physically present and used in a MME
;   it is explicitly optional for an operator to provision in this release of 3GPP
;.  So it too is commented out

; NAPTR order plays no major role in this particular example since the MME node is already
; selected in GUTI case and as a canonical node name. In most cases the interface type
;(S10 vs S11 etc) is functionally determined so the NAPTR order is rarely used in this record set
; If the S3 and Gn records were not commented out the S10 is preferred over S3 over Gn
; I.e. a combined MME/SGSN  could communicate to the MME above using any of the three protocols
```

```
; at context transfer.
; So the operator is stating that S10 is preferred over S3 over Gn
;
; Of course if the MME had multiple S10 interfaces
; the operator could provision more than one S10 record with different orders
; perhaps to select S10 IPv6 over S10 IPv4


; We have the same type of records for the other MME (same comments would apply)

;
mmec02.mmegi8001.mme                        (
;     IN NAPTR order pref. flag service                    regexp  replacement
      IN NAPTR 100   999   "a"   "x-3gpp-mme:x-s10"           ""  topoff.eth1.mmec02.mmegi8001.mme )
      IN NAPTR 200   999   "a"   "x-3gpp-mme:x-s11"           ""  topoff.eth3.mmec02.mmegi8001.mme
;     IN NAPTR 300   999   "a"   "x-3gpp-mme:x-s3"            ""  topoff.eth5.mmec02.mmegi8001.mme
;     IN NAPTR 400   999   "a"   "x-3gpp-mme:x-gn"            ""  topoff.eth6.mmec02.mmegi8001.mme
;     IN NAPTR 500   999   "a"   "x-3gpp-mme:x-s1-mme"        ""  topoff.eth7.mmec02.mmegi8001.mme
;
;
; All MME IP addresses for both MME
;
topoff.eth1.mmec01.mmegi8001.mme IN A 192.0.2.11
                                 IN A 192.0.2.12
                                 IN AAAA 2001:db8:0:0:0:0:0:0
                                 IN AAAA 2001:db8:0:1:0:0:0:0
topoff.eth3.mmec01.mmegi8001.mme IN A 192.0.2.13
                                 IN A 192.0.2.14
                                 IN AAAA 2001:db8:0:2:0:0:0:0
                                 IN AAAA 2001:db8:0:3:0:0:0:0
topoff.eth1.mmec02.mmegi8001.mme IN A 192.0.2.17
                                 IN A 192.0.2.18
                                 IN AAAA 2001:db8:0:6:0:0:0:0
                                 IN AAAA 2001:db8:0:7:0:0:0:0
topoff.eth3.mmec02.mmegi8001.mme IN A 192.0.2.19
                                 IN A 192.0.2.110
                                 IN AAAA 2001:db8:0:8:0:0:0:0
                                 IN AAAA 2001:db8:0:9:0:0:0:0

; end of file
```

The partially qualified MME host names are topoff.eth1.mmec01.mmegi8001., topoff.eth3.mmec02.mmegi8001.mme and similar. The (partially qualified) canonical node names are mmec01.mmegi8001.mme and mmec02.mmegi8001.mme (obtained by stripping off the first two labels of the host names).

The fully qualified MME node names are the relatively long

mmec01.mmegi8001.mme.epc.mnc990.mcc311.3gppnetwork.org.

and

mmec02.mmegi8001.mme.epc.mnc990.mcc311.3gppnetwork.org.

which are obtained by appending the value of $ORIGIN to the partially qualified MME node names. As stated before we will use the partially qualified names in this Annex to avoid this visual clutter and for typographical reasons.

Note that hostnames such as topoff.eth1.mmec01.mmegi8001.mme intentionally give back the node name mmec01.mmegi8001.mme so this record set is both for GUTI lookup and MME canonical node name lookup (and it is self consistent).

For the purposes of GUTI lookup the S10 record is sufficient for LTE only access and LTE only interworking. As stated in the comments above, S3 would be added for S4-SGSN support and the Gn/Gp records for Gn/Gp SGSN support.

For the purposes of being a canonical node record the S10 and S11 MME records are included (and they are mandatory to provision in the canonical node record of the MME).

NOTE 1: The rule for canonical node records is to always include all services that are allowed to be used on a node and are defined in 3GPP TS 23.003 [4] subclause 19.4.3 for that node. Node records are meant to be complete. We have an exception here since x-3gpp-mme:x-s1-mme is explicitly optional in this release of 3GPP.

A.4.5 APN file for "Collocated Simple LTE Example"

Reminder, the format of the APN FQDN is standardized by 3GPP TS 23.003 [4] subclause 19.4.2.2 and is of form <APN-NI>.apn.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org

Service names are in 3GPP TS 23.003 [4] subclause 19.4.3

There are two APN-NI in this network "imsTV1", and "imsTV2".

The file containing the APN records for this example will be CS_APN_DB.txt and has following content.

```
;
;
imsTV1.apn                                  (
;  IN NAPTR order pref. flag service                          regexp replacement
   IN NAPTR 100   999   "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp"        "" topoff.vip1.gw01.nodes )
   IN NAPTR 200   999   "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp"        "" topoff.vip1.gw21.nodes
;  IN NAPTR 300   999   "a" "x-3gpp-pgw:x-gn:x-gp"                "" topoff.vip3.gw01.nodes
;  IN NAPTR 400   999   "a" "x-3gpp-pgw:x-gn:x-gp"                "" topoff.vip3.gw21.nodes
   IN NAPTR 500   999   "a" "x-3gpp-pgw:x-s8-pmip"               "" topoff.vip2.gw01.nodes
   IN NAPTR 600   999   "a" "x-3gpp-pgw:x-s8-pmip"               "" topoff.vip2.gw21.nodes

;
;  Operator has imsTV1.apn using gw01.nodes when possible. Possibly one IMS core and video server
;  are closer to gw01.nodes than gw21.nodes
;
;  Operator only allows PMIPv6 for roaming type scenarios and only as a last choice
;  The operator does not support Gn/Gp so those records are commented out


imsTV2.apn                                  (
;  IN NAPTR order pref. flag service                          regexp replacement
   IN NAPTR 200   999   "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp"        "" topoff.vip1.gw01.nodes )
   IN NAPTR 100   999   "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp"        "" topoff.vip1.gw21.nodes
;  IN NAPTR 400   999   "a" "x-3gpp-pgw:x-gn:x-gp"                "" topoff.vip3.gw01.nodes
;  IN NAPTR 300   999   "a" "x-3gpp-pgw:x-gn:x-gp"                "" topoff.vip3.gw21.nodes
   IN NAPTR 600   999   "a" "x-3gpp-pgw:x-s8-pmip"               "" topoff.vip2.gw01.nodes
   IN NAPTR 500   999   "a" "x-3gpp-pgw:x-s8-pmip"               "" topoff.vip2.gw21.nodes
;
;  This is almost the same as imsTV1.
;  However, NAPTR order values for a particular interface type are reverse in comparision
;  to imsTV1.apn
;
;  Operator has imsTV2.apn using gw21.nodes when possible.


; Obviously more APN would exist for a real operator

;
; end of file
```

If collocation were not a consideration then the APN-NI "imsTV1" causes topoff.vip1.gw01.nodes to be selected and the APN-NI "imsTV2" causes topoff.vip1.gw21.nodes to be selected. This is because in the corresponding record set the order 100 record is taken before the order 200 record in the S-NAPTR. In case of failure of a PGW then the other PGW can be used (the record with order 200). If there is deliberate bias to a specific PGW is solely the concern of this operator, the S-NAPTR procedure simply honors this.

However, collocation is a consideration (topological checks are not relevant here since all records use topoff).

The (partially qualified) PGW node names are gw01.nodes and gw21.nodes (obtained by stripping off the first two labels of the partially qualified host names topoff.vip1.gw01.nodes, topoff.vip1.gw21.nodes and so on).

A.4.6 PGW/SGW node file for "Collocated Simple LTE Example"

The file containing the PGW/SGW canonical node name records (and A/AAAA records for the PGW/SGW nodes) for this example will be CS_SGW_PGW_NODE_DB.txt and has following content.

```
; Node records for the combined PGW/SGW


gw01.nodes                              (
;  IN NAPTR order pref. flag service                              regexp replacement
   IN NAPTR 200   999   "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp"          "" topoff.vip1.gw01.nodes )
   IN NAPTR 400   999   "a" "x-3gpp-pgw:x-s8-pmip"                 "" topoff.vip2.gw01.nodes
;  IN NAPTR 500   999   "a" "x-3gpp-pgw:x-gn:x-gp"                 "" topoff.vip3.gw01.nodes
;  Above record is commented out since this operator does not support Release 8 Gn/Gp functions
;  Would have to be included otherwise
;

; Above are PGW records.
; Note this operator does NOT support PMIP for S5
; Following are SGW records
   IN NAPTR 100   999   "a" "x-3gpp-sgw:x-s11"                     "" topoff.eth1.gw01.nodes
   IN NAPTR 300   999   "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp"         "" topoff.eth4.gw01.nodes
   IN NAPTR 600   999   "a" "x-3gpp-sgw:x-s8-pmip"                 "" topoff.eth9.gw01.nodes
;  IN NAPTR 700   999   "a" "x-3gpp-sgw:x-s4"                      "" topoff.eth6.gw01.nodes
;  IN NAPTR 710   999   "a" "x-3gpp-sgw:x-s12"                     "" topoff.eth6.gw01.nodes
;  IN NAPTR 800   999   "a" "x-3gpp-sgw:x-gn:x-gp"                 "" topoff.eth8.gw01.nodes
;  Above records are commented out since this operator does not support any SGSN variant
;  or UTRAN support. The operator is pure LTE in this example
;  Above records would be included if the SGWs were able to be used with the corresponding
;  SGSN function (or direct tunnel function)
;
;  Reminder: Canonical node records must be complete.
;
;   Exception, "x-3gpp-sgw:x-s1-u" records could be listed but are purely optional for an operator
to provision
;    in this release of 3GPP.

; NAPTR order plays no real role in this particular example (except for S8)
; Reasons for this example are:
; This is the PGW/SGW canonical node record so there is no node selection based from the record
; The interface type is functionally determined in most use cases using this record set except S5/S8
; This operator does NOT support PMIP for S5 and there is only one S5-GTP record.
; So order is not important except for S8.
; For S8 the operator does places PMIP with highest order just to be sure that GTP based S8 will be
; used first when possible even at re-selection of an S8 interface on the SGW or PGW as per this
; particular operators policy
; This is of course subject to the roaming agreements this operator has.



; Same record for the other combined PGW/SGW


gw21.nodes                              (
;  IN NAPTR order pref. flag service                              regexp replacement
   IN NAPTR 200   999   "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp"          "" topoff.vip1.gw21.nodes )
   IN NAPTR 400   999   "a" "x-3gpp-pgw:x-s8-pmip"                 "" topoff.vip2.gw21.nodes
;  IN NAPTR 500   999   "a" "x-3gpp-pgw:x-gn:x-gp"                 "" topoff.vip3.gw21.nodes

; Above are PGW records.

; Following are SGW records
   IN NAPTR 100   999   "a" "x-3gpp-sgw:x-s11"                     "" topoff.eth1.gw21.nodes
   IN NAPTR 300   999   "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp"         "" topoff.eth4.gw21.nodes
   IN NAPTR 600   999   "a" "x-3gpp-sgw:x-s8-pmip"                 "" topoff.eth9.gw21.nodes
;  IN NAPTR 700   999   "a" "x-3gpp-sgw:x-s4"                      "" topoff.eth6.gw21.nodes
;  IN NAPTR 710   999   "a" "x-3gpp-sgw:x-s12"                     "" topoff.eth6.gw21.nodes
;  IN NAPTR 800   999   "a" "x-3gpp-sgw:x-gn:x-gp"                 "" topoff.eth8.gw21.nodes



; A/AAAA records
;
; IP addresses for gw01
;
topoff.vip1.gw01.nodes IN A 192.0.2.113
                       IN A 192.0.2.114
                       IN AAAA 2001:db8:0:c:0:0:0:0
                       IN AAAA 2001:db8:0:d:0:0:0:0
topoff.vip2.gw01.nodes IN A 192.0.2.143
                       IN A 192.0.2.144
                       IN AAAA 2001:db8:0:2a:0:0:0:0
```

```
                        IN AAAA 2001:db8:0:2b:0:0:0:0
topoff.eth1.gw01.nodes IN A 192.0.2.129
                        IN A 192.0.2.130
                        IN AAAA 2001:db8:0:1c:0:0:0:0
                        IN AAAA 2001:db8:0:1d:0:0:0:0
topoff.eth4.gw01.nodes IN A 192.0.2.131
                        IN A 192.0.2.132
                        IN AAAA 2001:db8:0:1e:0:0:0:0
                        IN AAAA 2001:db8:0:1f:0:0:0:0
topoff.eth9.gw01.nodes IN A 192.0.2.133
                        IN A 192.0.2.134
                        IN AAAA 2001:db8:0:20:0:0:0:0
                        IN AAAA 2001:db8:0:21:0:0:0:0

; IP addresses for gw21
;
topoff.vip1.gw21.nodes IN A 192.0.2.115
                        IN A 192.0.2.116
                        IN AAAA 2001:db8:0:e:0:0:0:0
                        IN AAAA 2001:db8:0:f:0:0:0:0
topoff.vip2.gw21.nodes IN A 192.0.2.135
                        IN A 192.0.2.136
                        IN AAAA 2001:db8:0:22:0:0:0:0
                        IN AAAA 2001:db8:0:23:0:0:0:0
topoff.eth1.gw21.nodes IN A 192.0.2.137
                        IN A 192.0.2.138
                        IN AAAA 2001:db8:0:24:0:0:0:0
                        IN AAAA 2001:db8:0:25:0:0:0:0
topoff.eth4.gw21.nodes IN A 192.0.2.139
                        IN A 192.0.2.140
                        IN AAAA 2001:db8:0:26:0:0:0:0
                        IN AAAA 2001:db8:0:27:0:0:0:0
topoff.eth9.gw21.nodes IN A 192.0.2.141
                        IN A 192.0.2.142
                        IN AAAA 2001:db8:0:28:0:0:0:0
                        IN AAAA 2001:db8:0:29:0:0:0:0
;end of file
```

The above records are essentially standalone records representing the PGW/SGW node and its interfaces. An operator would normally define the records when a new node is brought into service and would only modify it when new IP interfaces or services are added to that particular node.

The fact both record types are here under the same name is a result of the PGW/SGW being collocated and having the same canonical node name.

NOTE: The rule for node records is to always include all services that are allowed to be used on a node and are defined in 3GPP TS 23.003 [4] subclause 19.4.3 for that node. Node records are meant to be complete.

The lack of a record in a canonical node name record set is just as important as the records that are there. The commented out records in the example would be included for supporting S4-SGSN and/or Gn/Gp interfaces. Their absence states that the node does not support the function.

A.4.7 TAI/TAC file for "Collocated Simple LTE Example"

The format of the TAI/TAC name is standardized by 3GPP 23.003 [4] subclause 19.4.2.3 and is of form tac-lb<TAC-low-byte>.tac-hb<TAC-high-byte>.tac.epc.mnc<MNC>.mcc<MCC>.3gppnetwork.org and service names are in 3GPP 23.003 [4] subclause 19.4.3.

The file containing the TAI/TAC records for this example will be SIMPLE_TAI_DB.txt and has following content.

```
; All TAC codes for one region

*.tac-hb01.tac                              (
;  IN NAPTR order pref. flag service                        regexp replacement
   IN NAPTR 100   999   "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp"      ""  topoff.eth4.gw01.nodes )
   IN NAPTR 200   999   "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp"      ""  topoff.eth4.gw21.nodes
   IN NAPTR 300   999   "a" "x-3gpp-sgw:x-s8-pmip"              ""  topoff.eth9.gw01.nodes
   IN NAPTR 400   999   "a" "x-3gpp-sgw:x-s8-pmip"              ""  topoff.eth9.gw21.nodes
; Above records are needed for SGW selection in initial attach of a UE (or TAU or handover attach)
;
   IN NAPTR 500   999   "a" "x-3gpp-mme:x-s10"                  ""  topoff.eth1.mmec01.mmegi8001.mme
   IN NAPTR 600   999   "a" "x-3gpp-mme:x-s10"                  ""  topoff.eth1.mmec02.mmegi8001.mme
```

```
;  Above two records are needed for target MME selection by source MME
;
;  IN NAPTR 700   999   "a" "x-3gpp-sgw:x-s11"                  ""  topoff.eth1.gw01.nodes )
;  IN NAPTR 800   999   "a" "x-3gpp-sgw:x-s11"                  ""  topoff.eth1.gw21.nodes
;  Above two S11 records are purely optional for an operator to provision and are only
;  an optimizaton when included so they are commented out for this example
;  IN NAPTR 900   999   "a" "x-3gpp-sgw:x-s4"                   ""  topoff.eth6.gw01.nodes )
;  IN NAPTR 1000  999   "a" "x-3gpp-sgw:x-s4"                   ""  topoff.eth6.gw21.nodes
; This operator does not support S3/S4 so they are commented out for this example
; Above two S4 records are purely optional for an operator to provision even if S3/S4 is supported

; Note relative value of NAPTR order is important between the S5/S8 records,
; relative value of NAPTR order is important between the S10 records,
; relative value of NAPTR order is important between the S11 records
; but is not really important between different interface types used here
; (i.e. the MME selection procedure does not look for an SGW interface)
;
;  This operators policy is PMIPv6 is used only as last choice
;  and only for S8 they don't allow S5 PMIPv6 at all


; All TAC codes for another region

*.tac-hb40.tac                          (
;  IN NAPTR order pref. flag service                        regexp replacement
   IN NAPTR 200   999   "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp"      ""  topoff.eth4.gw01.nodes )
   IN NAPTR 100   999   "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp"      ""  topoff.eth4.gw21.nodes
   IN NAPTR 400   999   "a" "x-3gpp-sgw:x-s8-pmip"              ""  topoff.eth9.gw01.nodes
   IN NAPTR 300   999   "a" "x-3gpp-sgw:x-s8-pmip"              ""  topoff.eth9.gw21.nodes
;
   IN NAPTR 600   999   "a" "x-3gpp-mme:x-s10"                  ""  topoff.eth1.mmec01.mmegi8001.mme
   IN NAPTR 500   999   "a" "x-3gpp-mme:x-s10"                  ""  topoff.eth1.mmec02.mmegi8001.mme
;
;  IN NAPTR 800   999   "a" "x-3gpp-sgw:x-s11"                  ""  topoff.eth1.gw01.nodes )
;  IN NAPTR 700   999   "a" "x-3gpp-sgw:x-s11"                  ""  topoff.eth1.gw21.nodes

;  IN NAPTR 1000  999   "a" "x-3gpp-sgw:x-s4"                   ""  topoff.eth6.gw01.nodes )
;  IN NAPTR 900   999   "a" "x-3gpp-sgw:x-s4"                   ""  topoff.eth6.gw21.nodes


; For the example the TAC values *.tac-hb40.tac are on other side of network than the *.tac-hb01.tac
; Relative order reverses in comparison to since those TAI are closer to the other set of nodes

; end of file
```

SGW records are here since TAI is used to select the SGW in the initial attach procedure (and TAU with SGW change). The MME record is here for the the 'Inter eNodeB handover with MME relocation' procedure where a target MME must be selected by the source MME.

Wild cards are used in this example to simplify provisioning assuming that the operator picks tac codes with the same high byte for each eNodeB that are equivalent in terms of prioritization. Otherwise each TAC code needs to be provisioned individually in DNS requiring a small amount of additional repetitive provisioning as each set of TAC codes is added to a network.

From above SGW records we see only GTP based S5 so there is no PMIPv6 S5 support for the SGWs. The operator insists on GTP based S5 i.e. when one of their PGW is used and the UE is in their network (the operator does let roamers use another operators PMIPv6 PGW though). The S5/S8 SGW interfaces with (partially qualified) canonical node name gw01.nodes is used when possible for tac codes with high byte of 01 hex and the S5/S8 SGW interfaces with (partially qualified) canonical node name gw21.nodes is used when possible for tac codes with high byte of 40 hex. This follows from the NAPTR order 100 record being used before the order 200 record in the corresponding record set. This is the general mechanism an operator can use to steer the SGW selection based on which TAI the user initially attaches to (or during TAU with SGW service change). So an operator trying to minimize S1-U link costs would set the order to the smallest value for the SGW with the least cost S1-U link(s) to the eNodeB(s) serving that TAI and use higher order for the TAI corresponding to the more expensive links.

The gw01.nodes and gw21.nodes are the collocated PGW/SGW canonical node names the operator chooses. These were defined in the file CS_SGW_PGW_NODE_DB.txt.

The TAI records are not intended to determine the SGW S11 interface directly but only the candidate list SGW S5/S8 interfaces with node names. S11 interface lookup can be done from the SGW canonical node name (see records in CS_SGW_PGW_NODE_DB.txt) so inclusion of S11 here is not strictly needed.

A.4.8 MME lookup based on GUTI for "Simple LTE Example"

Here we "manually" execute the S-NAPTR procedure using the standard "dig" command for finding the source MME at the target MME (for example in a TAU procedure) to illustrate the procedure and show how the DNS records work in practice.

The target MME receives the LTE UE's old GUTI. The target MME extracts the MNC, MCC, MMEC and MMEGI values. In this example the UE is in a cell where the MME selected by the RAN is assumed to be MMEC=1 and MMEGI=8000 hexadecimal. The MME creates creates the Application Unique String mmec01.mmegi8001.mme.$ORIGIN (see subclause 4.3.3.4 of this TS)

NOTE: Reminder $ORIGIN is epc.mnc990.mcc311.3gppnetwork.org. and is employed here simply to keep the length of the example text manageable.

The target MME starts the S-NAPTR with Application Unique String = mmec01.mmegi8001.mme.$ORIGIN and desired services x-3gpp-mme:x-s10 since it is trying to contact the old MME to do a context transfer.

Here we emulate the same action manually with the dig command.

```
--------- Command to DNS server----------
command: dig @192.0.2.247  +tcp  NAPTR mmec01.mmegi8001.mme.$ORIGIN
---------Start Response from DNS server----------

; <<>> DiG 9.5.0-P2-W2 <<>> @192.0.2.247 +tcp NAPTR mmec01.mmegi8001.mme.$ORIGIN
; (1 server found)
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 1399
;; flags: qr aa rd; QUERY: 1, ANSWER: 2, AUTHORITY: 2, ADDITIONAL: 10
;; WARNING: recursion requested but not available

;; QUESTION SECTION:
;mmec01.mmegi8001.mme.$ORIGIN.  IN NAPTR

;; ANSWER SECTION:
mmec01.mmegi8001.mme.$ORIGIN.   3600 IN NAPTR 100 999 "a" "x-3gpp-mme:x-s10" ""
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN.
mmec01.mmegi8001.mme.$ORIGIN.   3600 IN NAPTR 200 999 "a" "x-3gpp-mme:x-s11" ""
topoff.eth3.mmec01.mmegi8001.mme.$ORIGIN.

;; AUTHORITY SECTION:
$ORIGIN. 3600   IN NS   ns1.$ORIGIN.
$ORIGIN. 3600   IN NS   ns2.$ORIGIN.

;; ADDITIONAL SECTION:
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.12
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.11
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8::
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:1::
topoff.eth3.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.14
topoff.eth3.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.13
topoff.eth3.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:2::
topoff.eth3.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:3::
ns1.$ORIGIN. 3600 IN A 192.0.2.247
ns2.$ORIGIN. 3600 IN A 192.0.2.248

;; Query time: 15 msec
;; SERVER: 192.0.2.247#53(192.0.2.247)
;; WHEN: Wed Jan 21 16:30:20 2009
;; MSG SIZE  rcvd: 524

---------End Response from DNS server----------
```

MME retains only NAPTR with matching services matching x-3gpp-mme:x-s10 yielding

```
NAPTR record set
replacement                              service                       flag order preference
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN x-3gpp-mme:x-s10              "a"  100   999
```

MME node sorts the NAPTR records by RFC 3958 rules. Obviously since this is a list of only one record it yields the same list back again unchanged.

Since the record is the terminal ("a") flag we have the host name (and canonical node name inside it).

Formally the A/AAAA record lookup can be performed at this point or it can be deferred until a later step.

The MME can store the important content now

topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN services of x-3gpp-mme:x-s10

Because there is only one match the MME now has the final candidate list (except for IP addresses).

topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN services of x-3gpp-mme:x-s10

Note the needed A/AAAA records for topoff.eth1.mmec01.mmegi8001.mme are already present in the additional record section of the response.

```
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.12
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.11
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8::
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:1::
```

If the records had not been available in the additional record section of the response then the MME would have had to done the A/AAAA record lookups directly. Emulated by dig command they would have been

```
command: dig @192.0.2.247  +tcp  A topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN
```
and

```
command: dig @192.0.2.247  +tcp  AAAA topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN
```

or

```
command: dig @192.0.2.247  +tcp  ANY topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN
```

Regardless of the source of the A and AAAA records, the order of the A records shall be randomly shuffled by the MME. The same shall be done for the AAAA records. (For our "manual" example it will be a coin flip as to which of the two records is first).

The MME now has a candidate list (very short here) with the needed IPv4 and IPv6 addresses. I.e.:

(topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN, services of x-3gpp-mme:x-s10, {192.0.2.12,192.0.2.11}, { 2001:db8::, 2001:db8:0:1:: })

The target MME then can use the IPv4 (or IPv6) addresses to contact the source MME over S10 to get the UE's context information from the old MME.

When an operator uses only NAPTR flag "a" records the S-NAPTR procedure will only require one NAPTR lookup from the Application Unique String in all cases greatly streamlining the DNS lookups. Also assuming a modern DNS server where the NAPTR and A/AAAA records in the same DNS server the needed A/AAAA records are likely to be in the additional record section (as the above real output shows).

However, the needed A/AAAA records would NOT have been present even in this small example if the DNS resolver (dig here) had not used TCP transport (or EDNS0 for UDP) since the packet size of the response in this example was 524 bytes which exceeds the 512 byte limit for using UDP transport without EDNS0.

A.4.9 APN lookup for "Simple LTE Example" (i.e. PGW candidate list)

Assume a non-roaming LTE UE indicates it want to use APN-NI string "imsTV2" to the MME in our example network at initial attach.

NOTE 1: Reminder $ORIGIN is epc.mnc990.mcc311.3gppnetwork.org. and is employed here simply to keep the length of the example text manageable

The MME starts the S-NAPTR procedure with Application Unique String = imsTV2.apn.$ORIGIN and desired services x-3gpp-pgw:x-s5-gtp and x-3gpp-pgw:x-s5-pmip.

Here we emulate the same action the MME would do manually with the dig command.

MME starts with Application Unique String = imsTV2.apn.$ORIGIN and the desired services x-3gpp-pgw:x-s5-gtp and x-3gpp-pgw:x-s5-pmip

```
--------- Command to DNS server----------
command: dig @192.0.2.247  +tcp  NAPTR imsTV2.apn.$ORIGIN
---------Start Response from DNS server----------

; <<>> DiG 9.5.0-P2-W2 <<>> @192.0.2.247 +tcp NAPTR imsTV2.apn.$ORIGIN
; (1 server found)
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 1796
;; flags: qr aa rd; QUERY: 1, ANSWER: 4, AUTHORITY: 2, ADDITIONAL: 18
;; WARNING: recursion requested but not available

;; QUESTION SECTION:
;imsTV2.apn.$ORIGIN. IN NAPTR

;; ANSWER SECTION:
imsTV2.apn.$ORIGIN. 3600 IN NAPTR 600 999 "a" "x-3gpp-pgw:x-s8-pmip" ""
topoff.vip2.gw01.nodes.$ORIGIN.
imsTV2.apn.$ORIGIN. 3600 IN NAPTR 100 999 "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp" ""
topoff.vip1.gw21.nodes.$ORIGIN.
imsTV2.apn.$ORIGIN. 3600 IN NAPTR 200 999 "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp" ""
topoff.vip1.gw01.nodes.$ORIGIN.
imsTV2.apn.$ORIGIN. 3600 IN NAPTR 500 999 "a" "x-3gpp-pgw:x-s8-pmip" ""
topoff.vip2.gw21.nodes.$ORIGIN.

;; AUTHORITY SECTION:
$ORIGIN. 3600   IN NS   ns2.$ORIGIN.
$ORIGIN. 3600   IN NS   ns1.$ORIGIN.

;; ADDITIONAL SECTION:
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.116
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.115
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:e::
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:f::
topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.114
topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.113
topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:c::
topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:d::
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.136
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.135
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:22::
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:23::
topoff.vip2.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.144
topoff.vip2.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.143
topoff.vip2.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:2a::
topoff.vip2.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:2b::
ns1.$ORIGIN. 3600 IN A 192.0.2.247
ns2.$ORIGIN. 3600 IN A 192.0.2.248

;; Query time: 0 msec
;; SERVER: 192.0.2.247#53(192.0.2.247)
;; WHEN: Wed Jan 21 19:21:14 2009
;; MSG SIZE  rcvd: 890

---------End Response from DNS server----------
MME retains only NAPTR records with matching services x-3gpp-pgw:x-s5-gtp and
x-3gpp-pgw:x-s5-pmip  yielding:
NAPTR record set
replacement                    service                         flag order preference
topoff.vip1.gw21.nodes.$ORIGIN x-3gpp-pgw:x-s5-gtp:x-s8-gtp   "a"  100   999
topoff.vip1.gw01.nodes.$ORIGIN x-3gpp-pgw:x-s5-gtp:x-s8-gtp   "a"  200   999
; Note the x-s8-gtp are not really included but are kept here to allow the
; reader to see which NAPTR record was kept from the DNS response.
; The DNS server "luckily" returns this in sorted order but the MME must sort it .
; The DNS server is not responsible for sorting but the DNS resolver (i.e. the MME)
MME node sorts NAPTR by RFC 3958 yielding
NAPTR record set
replacement                    service                         flag order preference
topoff.vip1.gw21.nodes.$ORIGIN x-3gpp-pgw:x-s5-gtp:x-s8-gtp   "a"  100   999
topoff.vip1.gw01.nodes.$ORIGIN x-3gpp-pgw:x-s5-gtp:x-s8-gtp   "a"  200   999
MME Stores records since they are flag "a"
topoff.vip1.gw21.nodes.$ORIGIN services of x-3gpp-pgw:x-s5-gtp
```

```
topoff.vip1.gw01.nodes.$ORIGIN services of x-3gpp-pgw:x-s5-gtp
MME now has final candidate list (A and AAAA lookup is deferred for our hand example)
topoff.vip1.gw21.nodes.$ORIGIN services of x-3gpp-pgw:x-s5-gtp
topoff.vip1.gw01.nodes.$ORIGIN services of x-3gpp-pgw:x-s5-gtp
```

The needed A/AAAA records were included with additional records. I.e.

```
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.116
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.115
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:e::
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:f::

and

topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.114
topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.113
topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:c::
topoff.vip1.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:d::
```

IF the A and AAAA records were not available in the Additional Record section (or a DNS cache) the MME would do the A/AAAA lookups. Which emulating with manual commands would look like:

```
dig @192.0.2.247  +tcp  A topoff.vip1.gw21.nodes.$ORIGIN

dig @192.0.2.247  +tcp  AAAA topoff.vip1.gw21.nodes.$ORIGIN

dig @192.0.2.247  +tcp  A topoff.vip1.gw01.nodes.$ORIGIN

dig @192.0.2.247  +tcp  AAAA topoff.vip1.gw01.nodes.$ORIGIN
```

We can form the full candidate list now (after random shuffling the A and AAAA records) to get

```
(topoff.vip1.gw21.nodes.$ORIGIN ,services of x-3gpp-pgw:x-s5-gtp , {192.0.2.115,192.0.2.116}, {
2001:db8:0:e::,2001:db8:0:f::} )
(topoff.vip1.gw01.nodes.$ORIGIN ,services of x-3gpp-pgw:x-s5-gtp , {192.0.2.114,192.0.2.113}, {
2001:db8:0:c::, 2001:db8:0:d::)
```

The fact that the node topoff.vip1.gw21.nodes.$ORIGIN comes first is a direct result of the two different NAPTR order values.

NOTE 2: This is the last example in this Annex with A and AAAA lookups and IP addresses explicitly included in the candidate list. This detail is left for the reader for the remaining sections.

For the initial attach case the candidate list of SGW has to be formally found to handle the collocated PGW and SGW case. The SGW candidate list is found in the next section.

NOTE 3: If this DNS procedure was being used for an additional PDN connection for a UE with an existing SGW rather than a first attach, then the MME would check to see if the current SGW node name matches any of the PGW node names in the above list (gw01.nodes.$ORIGIN and gw21.nodes.$ORIGIN) to try to pick a colocated PGW on the current SGW. The canonical node name record of the current SGW can be used by the MME to find all S5/S8 SGW interfaces on the current PGW. This is a smaller example of what follows so we do not detail it further.

A.4.10 TAI lookup for "Simple LTE Example" (i.e. SGW candidate list)

Assume a non-roaming LTE UE performs and initial attach and indicates it want to use a APN-NI (an APN in this operator's network). The MME knows it does not need to use S8 since it is a non-roaming UE and must be local APN so S5 is used.

NOTE 1: Reminder $ORIGIN is epc.mnc990.mcc311.3gppnetwork.org. and is employed here simply to keep the length of the example text manageable.

The MME has the TAI value where the UE has attached. We assume the low byte of the TAC is hex 11 and the high byte is hex 40.

The MME starts the S-NAPTR procedure with Application Unique String = tac-lb11.tac-hb40.tac.$ORIGIN and desired services x-3gpp-sgw:x-s11,x-3gpp-sgw:x-s5-gtp, x-3gpp-sgw:x-s5-pmip.

NOTE 2: This particular MME vendor looks for the x-s11 values, which is only an optional optimization. This will not have any benefit in this example since the operator chosed not to provision them.

Here we emulate the same action the MME would do manually with the "dig" command.

```
MME starts with Application Unique String = tac-lb11.tac-hb40.tac.$ORIGIN
desired services x-3gpp-sgw:x-s11,x-3gpp-sgw:x-s5-gtp,x-3gpp-sgw:x-s5-pmip

--------- Command to DNS server----------
command: dig @192.0.2.247  +tcp  NAPTR tac-lb11.tac-hb40.tac.$ORIGIN
---------Start Response from DNS server----------

; <<>> DiG 9.5.0-P2-W2 <<>> @192.0.2.247 +tcp NAPTR tac-lb11.tac-hb40.tac.$ORIGIN
; (1 server found)
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 622
;; flags: qr aa rd; QUERY: 1, ANSWER: 6, AUTHORITY: 2, ADDITIONAL: 26
;; WARNING: recursion requested but not available

;; QUESTION SECTION:
;tac-lb11.tac-hb40.tac.$ORIGIN. IN NAPTR

;; ANSWER SECTION:
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 400 999 "a" "x-3gpp-sgw:x-s8-pmip" ""
topoff.eth9.gw01.nodes.$ORIGIN.
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 500 999 "a" "x-3gpp-mme:x-s10" ""
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN.
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 600 999 "a" "x-3gpp-mme:x-s10" ""
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN.
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 100 999 "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp" ""
topoff.eth4.gw21.nodes.$ORIGIN.
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 200 999 "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp" ""
topoff.eth4.gw01.nodes.$ORIGIN.
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 300 999 "a" "x-3gpp-sgw:x-s8-pmip" ""
topoff.eth9.gw21.nodes.$ORIGIN.

;; AUTHORITY SECTION:
$ORIGIN. 3600   IN NS   ns2.$ORIGIN.
$ORIGIN. 3600   IN NS   ns1.$ORIGIN.

;; ADDITIONAL SECTION:
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.140
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.139
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:26::
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:27::
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.132
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.131
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:1e::
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:1f::
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.142
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.141
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:28::
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:29::
topoff.eth9.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.134
topoff.eth9.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.133
topoff.eth9.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:20::
topoff.eth9.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:21::
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.18
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.17
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:6::
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:7::
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.12
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.11
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8::
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:1::
ns1.$ORIGIN. 3600 IN A 192.0.2.247
ns2.$ORIGIN. 3600 IN A 192.0.2.248

;; Query time: 0 msec
;; SERVER: 192.0.2.247#53(192.0.2.247)
;; WHEN: Wed Jan 21 19:33:00 2009
```

```
;; MSG SIZE  rcvd: 1285

---------End Response from DNS server----------
MME retains only the NAPTR with matching services x-3gpp-sgw:x-s11,x-3gpp-sgw:x-s5-gtp,
and x-3gpp-sgw:x-s5-pmip yielding
NAPTR record set
replacement                    service                        flag order preference
topoff.eth4.gw01.nodes.$ORIGIN x-3gpp-sgw:x-s5-gtp:x-s8-gtp   "a"  200   999
topoff.eth4.gw21.nodes.$ORIGIN x-3gpp-sgw:x-s5-gtp:x-s8-gtp   "a"  100   999
; Note the x-s8-gtp are not really included but are kept here to allow the
; reader to see which NAPTR record was kept from the DNS response.

MME node sorts the NAPTR records by RFC 3958 yielding
NAPTR record set
replacement                  service                        flag order preference
topoff.eth4.gw21.nodes.$ORIGIN x-3gpp-sgw:x-s5-gtp:x-s8-gtp   "a"  100   999
topoff.eth4.gw01.nodes.$ORIGIN x-3gpp-sgw:x-s5-gtp:x-s8-gtp   "a"  200   999

MME Stores
topoff.eth4.gw21.nodes.$ORIGIN services of x-3gpp-sgw:x-s5-gtp
topoff.eth4.gw01.nodes.$ORIGIN services of x-3gpp-sgw:x-s5-gtp

MME now has candidate list
topoff.eth4.gw21.nodes.$ORIGIN services of x-3gpp-sgw:x-s5-gtp
topoff.eth4.gw01.nodes.$ORIGIN services of x-3gpp-sgw:x-s5-gtp


Again the A/AAAA records were available in the additional record section.

topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.140
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.139
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:26::
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:27::
and
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.132
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN A 192.0.2.131
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:1e::
topoff.eth4.gw01.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:1f::
```

NOTE: The size of the response was 1285 bytes and indicates why EDNS0 or TCP transport should be used for S-NAPTR when feasible. Most of the data in the Additional Record Section would not have been available if only 512 bytes had been sent.

A.4.11 Finding the collocated SGW and PGW together

IP addresses are not included here for readability.

We are now in a position to look at both the SGW candidate list and PGW candidate list at the same time.

From A.4.9 we had the PGW candidate list for the APN-NI "imsTV2" was

```
topoff.vip1.gw21.nodes.$ORIGIN services of x-3gpp-pgw:x-s5-gtp
topoff.vip1.gw01.nodes.$ORIGIN services of x-3gpp-pgw:x-s5-gtp
```

From A.4.10 we had the SGW candidate list for TAI where low byte ofTAC is hex 11 and the high byte is hex 40 was

```
topoff.eth4.gw21.nodes.$ORIGIN services of x-3gpp-sgw:x-s5-gtp
topoff.eth4.gw01.nodes.$ORIGIN services of x-3gpp-sgw:x-s5-gtp
```

We only have GTPv2 so the protocols match.

Both nodes gw21.nodes.$ORIGIN and gw01.nodes.$ORIGIN are collocated.

Hence, moving the collocated nodes to the front of both lists leaves them unchanged.

The procedure requires we use the SGW ordering to select the SGW first.

Hence, the SGW interface to try first is topoff.eth4.gw21.nodes.$ORIGIN

That is node gw21.nodes.$ORIGIN.

That means the PGW record that is collocated is:

```
topoff.vip1.gw21.nodes.$ORIGIN services of x-3gpp-pgw:x-s5-gtp
```
which is the first PGW interface to try.

NOTE: It is an "accident" that this is the same order as the APN-NI "imsTV2" gave back. If APN-NI had been "imsTV1" the order would have been reversed in the PGW candidate list and it would NOT have been first. Again the collocation ordering takes precedence.

The MME would use the IP addresses from the records (not shown) for the S11 Create Session Request attempt.

Assuming it succeeds we would not need to try any other PGW or SGW records.

The MME needs the S11 interface so it can physically send the Create Session Request to the SGW. That is covered in the next section.

A.4.12 S11 lookup by SGW canonical node name

Assume an SGW has been picked with canonical node name of gw21.nodes.$ORIGIN

The MME starts with Application Unique String = gw21.nodes.$ORIGIN obtained from the SGW selection with.desired services x-3gpp-sgw:x-s11

```
--------- Command to DNS server----------
command: dig @192.0.2.247  +tcp  NAPTR gw21.nodes.$ORIGIN
---------Start Response from DNS server----------

; <<>> DiG 9.5.0-P2-W2 <<>> @192.0.2.247 +tcp NAPTR gw21.nodes.$ORIGIN
; (1 server found)
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 23
;; flags: qr aa rd; QUERY: 1, ANSWER: 5, AUTHORITY: 2, ADDITIONAL: 22
;; WARNING: recursion requested but not available

;; QUESTION SECTION:
;gw21.nodes.$ORIGIN. IN NAPTR

;; ANSWER SECTION:
gw21.nodes.$ORIGIN. 3600 IN NAPTR 600 999 "a" "x-3gpp-sgw:x-s8-pmip" ""
topoff.eth9.gw21.nodes.$ORIGIN.
gw21.nodes.$ORIGIN. 3600 IN NAPTR 100 999 "a" "x-3gpp-sgw:x-s11" "" topoff.eth1.gw21.nodes.$ORIGIN.
gw21.nodes.$ORIGIN. 3600 IN NAPTR 200 999 "a" "x-3gpp-pgw:x-s5-gtp:x-s8-gtp" ""
topoff.vip1.gw21.nodes.$ORIGIN.
gw21.nodes.$ORIGIN. 3600 IN NAPTR 300 999 "a" "x-3gpp-sgw:x-s5-gtp:x-s8-gtp" ""
topoff.eth4.gw21.nodes.$ORIGIN.
gw21.nodes.$ORIGIN. 3600 IN NAPTR 400 999 "a" "x-3gpp-pgw:x-s8-pmip" ""
topoff.vip2.gw21.nodes.$ORIGIN.

;; AUTHORITY SECTION:
$ORIGIN. 3600   IN NS   ns2.$ORIGIN.
$ORIGIN. 3600   IN NS   ns1.$ORIGIN.

;; ADDITIONAL SECTION:
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.138
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.137
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:24::
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:25::
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.116
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.115
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:e::
topoff.vip1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:f::
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.139
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.140
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:27::
topoff.eth4.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:26::
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.136
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.135
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:22::
topoff.vip2.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:23::
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.141
```

```
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.142
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:29::
topoff.eth9.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:28::
ns1.$ORIGIN. 3600 IN A 192.0.2.247
ns2.$ORIGIN. 3600 IN A 192.0.2.248

;; Query time: 0 msec
;; SERVER: 192.0.2.247#53(192.0.2.247)
;; WHEN: Wed Jan 21 19:45:36 2009
;; MSG SIZE  rcvd: 1072

---------End Response from DNS server----------
MME retains only the NAPTR records with matching service x-3gpp-sgw:x-s11 yielding
NAPTR record set
replacement                    service                         flag order preference
topoff.eth1.gw21.nodes.$ORIGIN x-3gpp-sgw:x-s11                "a"  100   999
MME node sorts NAPTR by RFC 3958 (trivially here) yielding
NAPTR record set
replacement                    service                         flag order preference
topoff.eth1.gw21.nodes.$ORIGIN x-3gpp-sgw:x-s11                "a"  100   999
MME Stores
topoff.eth1.gw21.nodes.$ORIGIN services of x-3gpp-sgw:x-s11
MME now has candidate list
topoff.eth1.gw21.nodes.$ORIGIN services of x-3gpp-sgw:x-s11
```

Again the needed A and AAAA records are in the additional record section

```
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.138
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN A 192.0.2.137
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:24::
topoff.eth1.gw21.nodes.$ORIGIN. 3600 IN AAAA 2001:db8:0:25::
```

A.4.12 TAI lookup for "Colocation Simple LTE Example" (i.e. MME candidate list)

This type of lookup is done for the "Inter eNodeB handover with MME relocation" procedure where the old MME selects the target MME in a new TAI.

The service is x-3gpp-mme:x-s10 again.

Since this also starts with the TAI NAPTR lookup we have the same type of record set as the SGW selection from TAI section (see A.4.10) and to not waste space the same TAC value will be used.

The only difference from that section starts with which NAPTR records match.

From the Answer section these are:

```
;; ANSWER SECTION:
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 500 999 "a" "x-3gpp-mme:x-s10" ""
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN.
tac-lb11.tac-hb40.tac.$ORIGIN. 3600 IN NAPTR 600 999 "a" "x-3gpp-mme:x-s10" ""
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN.
```

Which are already in order. The NAPTR order of 500 is less than NAPTR order of 600.

So S10 interfaces in order are:

topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN.

topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN.

and the needed A/AAAA records from the Additional section in the DNS response in section A.4.10 are:

```
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.18
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.17
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:6::
topoff.eth1.mmec02.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:7::

topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.12
```

```
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN A 192.0.2.11
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8::
topoff.eth1.mmec01.mmegi8001.mme.$ORIGIN. 3600 IN AAAA 2001:db8:0:1::
```

Annex B (Normative): DNS procedures clarifications

B.1 DNS RFC procedures general clarifications

This sub-clause clarifies DNS resolver use of the S-NAPTR procedures in EPC core network nodes.

NOTE: The only EPC core network nodes identified explicitly at this time that employ S-NAPTR procedures are the MME and Release 8 SGSN.

DNS resolvers in EPC core network nodes shall support recursive queries and responses over UDP transport as specified in IETF RFC 1035 [3]. The EPC core network nodes may assume the existence of a local caching DNS server (see GSMA PRD IR.67 [5]) and hence may not need to do iterative queries as specified in IETF RFC 1035 [3]. However, the final deployment decision of local caching DNS servers is up to the operators. It is recommended that the EPC core network nodes support DNS queries and responses over TCP transport up to the 65535 byte maximum. Support of IETF RFC 2671 [13] (EDNS0) is recommended, in order to allow DNS response packets sizes over 512 octets when using UDP transport.

It is recommended that resolvers in EPC core network nodes cache frequently used DNS queries in order to lower load on DNS infrastructure.

EPC core network nodes shall support SRV records as specified in IETF RFC 2782 [8]. However, in the 3GPP scope the ordering of SRV records of the same priority SHALL use the algorithm described in IETF RFC 2782 [8] page 4 instead of the the "SHOULD" requirement in the IETF RFC. This is a 3GPP specific requirement to strengthen the described algorithm and actually allow predictable behavior of the IETF RFC 2782 [8] based load balancing.

B.2 DNS procedures 3GPP clarifications on S-NAPTR

IETF RFC 3958 [9] S-NAPTR procedures are unmodified with an exception of the following clarifications on the topological closeness and multi-protocol support:

1) For topological closeness the "topon" label matching of subclause 4.3.2 of the present document takes precedence over NAPTR ordering but NAPTR ordering is still used when matching label lengths are equal. Therefore, a full list of "candidate" records is needed as sketched in Appendix A.2 of IETF RFC 3958 [9], which in turn requires "backtracking" as described by IETF RFC 3958 [9] section 2.2.4. When collocation is to be considered applicable in a procedure it takes precedence in ordering over both "topon" and NAPTR ordering regardless of the value of the value of the "topon|topoff" label.

2) IETF RFC 3958 [9] has an ambiguity for S-NAPTR with multiple protocols in last paragraph of section 2.2.5 "It MAY choose to run simultaneous DDDS resolutions for more than one protocol, in which case the requirements above apply for each protocol independently. That is, do not switch protocols mid- resolution." The term " simultaneous DDDS resolutions" and "apply for each protocol independently" are not defined and can have different meanings. To resolve that ambiguity in S-NAPTR, the present document formally defines "Service description meeting the client requirement" from IETF RFC 3402 [14] section 3.3 step 4 as a NAPTR record with one or more of the 3GPP desired service and protocol field pair(s) and such that all ancestor NAPTR records in the current path to this point also include the identified service and protocol in the DDDS procedure. The present document uses that as the definition of "simultaneous DDDS resolutions". See subclause C.1 for more practical information on this point.

3) Strict ordering within S-NAPTR is obtained from the NAPTR order value as required in IETF RFC 3402 [14] and IETF RFC 3958 [8]. The value of (65535- NAPTR preference) shall be used as a statistical weight the same way the SRV weight is used on page 4 of IETF RFC 2782 [8] for SRV records.

Items 1) , 2) and 3) impact the ordering of DNS records in which they are returned by the S-NAPTR procedure. Items 1) and 2) also involve areas where the IETF RFC 3958 [9] only provides a sketch of the procedures needed and implicitly relies on IETF RFC 3402 [14] for details. To clarify these points as well as to guide implementations informative pseudo-code is provided in subclauses C.1, C.2 and C.3. Item 3) allows an operators to load balance within

one NAPTR record set without employing SRV records giving a simpler DNS provisioning and potentially reducing the number of DNS queries.

Annex C (Informative): DNS Pseudo-Code

C.1 S-NAPTR procedure base pseudo-code

The primary purpose of sub-clause C.1 is to show practically any differences that are normatively documented in subclause B.2. The changes to the IETF RFC 3958 [9] pseudo-code make this much clearer and self-contained than the normative text from subclause B.2.

The pseudo-code immediately following is the pseudo-code from IETF RFC 3958 [9] Appendix A.1 modified to incorporate the clarifications from subclause B.2.

```
  target = [initial domain]
; Next line is changed from Appendix A.1 of RFC 3958
  usable-service-protocol-set  = [initial desired service and protocol pairs]

naptr-done = false

   while (not naptr-done)
    {
     NAPTR-RRset = [DNSlookup of NAPTR RRs for target]
; Next two lines are changed from Appendix A.1 of RFC 3958
     NAPTR weight := 65535 - NAPTR order for each RR
     [sort NAPTR by ORDER, and by PREF within each ORDER and by statistical weight within each PREF]
     rr-done = false
     cur-rr = [first NAPTR RR]


     while (not rr-done)
; Next three lines are changed from Appendix A.1 of RFC 3958
        compatable-service-protocol-set =[ [usable-service-protocol-set] set intersection with
                                           [SERVICE field of cur-rr] ]
        if ([compatable-service-protocol-set] is not empty)
           rr-done = true
           target= [REPLACEMENT target of NAPTR RR]
; Next line is changed from Appendix A.1 of RFC 3958
           usable-service-protocol-set  = [compatable-service-protocol-set]
        else
           cur-rr = [next rr in list]

        if (not empty [FLAG in cur-rr])
           naptr-done = true
    }


port = -1


   if ([FLAG in cur-rr is "S"])
    {
     SRV-RRset = [DNSlookup of SRV RRs for target]
; Next line is changed from Appendix A.1
[Sort SRV RRset using the algorithm described on page 4 of IETF RFC 2782 [8]]
     target = [target of first RR of SRV-RRset]
     port = [port in first RR of SRV-RRset]
    }


   ; now, whether it was an "S" or an "A" in the NAPTR, we
   ; have the target for an A record lookup
;  Remaining lines are changed from Appendix A.1 of RFC 3958
   ; or AAAA record lookup

   IPv4_hosts = [DNSlookup of A RRs for target]
   IPv6_hosts = [DNSlookup of AAAA RRs for target]
   randomized order of IPv4_hosts and IPv6_hosts
   hostname = [target]
   return (hostname, usable-service-protocol-set, IPv4_hosts, IPv6_hosts, port)
```

The significant differences in the above Pseudo-Code and the IETF RFC 3958 [9] Pseudo-Code are :

A) [Sort SRV RRset using the algorithm described on page 4 of IETF RFC 2782 [8]]
which was changed from
[sort SRV-RRset based on PREF]
The Pseudo-Code in IETF RFC 3958 [9] simply has an error. There isn't even a PREF in a SRV record. Again see page 4 of IETF RFC 2782 [8] for the proper procedure.
The NAPTR weight is defined to be 65535 - NAPTR preference and is handled analogously to the SRV case.

B) IETF RFC 3958 [9] Appendix A.1 starts with "Assuming the client supports 1 protocol for a particular application" so the pseudo-code obviously was designed for one protocol at a time. The lines with usable-service-protocol-set and compatable-service-protocol-set above are the most important change to support multiple service/protocol combinations and are really the primary reason for providing the above Pseudo-Code.

There are two possible ways to interpret the last paragraph of section 2.2.5 of IETF RFC 3958 [9] when a list of multiple services/protocols is desired. One is the above interpretation using "set intersection" which allows multiple services/protocols. The other is to run the above procedure for one service and protocol at a time from the "desired service_and_protocol_set" and get a separate list for each service and protocol. In both approaches the relative ordering within a particular service and protocol is identical. If the proper interpretation of IETF RFC 3958 [9] is one service and protocol at a time, then the IETF RFC 3958 [9] does not define order between different service or protocols. Thus 3GPP is free to order between different 3GPP service and protocol types so long as the order within a service and protocol is respected. The above method does respect the order within a service and protocol therefore it is valid in either interpretation of section 2.2.5 of IETF RFC 3958 [9] and also valid in IETF RFC 3402 [14]).

The remaining changes in Pseudo-Code above are minor and mostly intended to show that the S-NAPTR procedure logically outputs following:

(hostname, usable-service-protocol-set, IPv4_hosts, IPv6_hosts, port)

where the returned hostname is the FQDN of the topologically aware node name with topon/topoff and interface information.

NOTE: Lookup of the A and AAAA records to get the IPv4 and IPv6 addresses may be deferred until they are needed to contact a selected server as an optional optimization.

In the 3GPP scope, a full implementation of RFC 3958 SHALL implement "backtracking" as described by IETF RFC 3958 [9] section 2.2.4 as required in subclause B.2.

For simplicity of the presentation in this Annex we assume a full IETF RFC 3958 [9] implementation with a call back interface as described in Appendix A.2 of the IETF RFC 3958 [9].

```
procedure S_NAPTR_to_callback(targetFQDN,
                              desired_service_and_protocol_set,
                              call_back_function)
```

where the call_back_function has interface

```
call_back_function (hostname, usable_service_and_protocol_set, port, IPv4_list,IP6_list)
```

The call_back_function returns "stop" if it does not want more records otherwise it returns "looking" and will be called with the next record.

C.2 S-NAPTR procedure - no topon

If topological closeness is not stated as specifically applicable to a procedure, or all node names are prefixed with "topoff", and if collocation is stated as not applicable to a procedure, then the first interface that can be successfully connected to would be sufficient to be returned from the S-NAPTR procedure. The following pseudo-code shows how the procedure works.

```
/*
 * The Callback function called from the S-NAPTR procedure
 * for each FQDN the S-NAPTR procedure finds..
 */
procedure try_to_connect (hostname, service_and_protocol_set, port, IP4_list, IP6_list)
  Begin procedure
//  Comment does procedure as outlined in C.1
    Use 3GPP procedures to try to connect in turn to all combinations
      of the service/protocols and IP addresses provided in the input.
    Upon first success return(stop);
    If all fail return(looking);
  End procedure;
};

/*
 * The S-NAPTR procedure follows
 *
 */

procedure connect_first_match (targetFQDN, desired_protocol_set)
  Begin procedure
    status:=S_NAPTR_to_callback(targetFQDN,
                                desired_service_and_protocol_set ,
                                try_to_connect);
    if status equals looking return(failure) else return(success);
  End procedure;
```

C.3 S-NAPTR procedure candidate list

The following procedure will get the complete candidate list. This is the "sorted list of matches" described in Appendix A.2 of IETF RFC 3958 [9]. This is used for an exhaustive search of all matches.

If the "topon" feature is specifically applicable to find "close" nodes, orcollocation is applicable, then the simple approach of getting the first match as described in subclause C.2 cannot be used. The S-NAPTR must be performed by exhaustive searching for all matching records since the best match by "topon" node name can be any record independent of S-NAPTR record ordering.

```
/*
 * The Callback function called by the S-NAPTR procedure for
 * each found match..
 */

procedure private_store_candidate_list (hostname,service_and_protocol_set, port,IP4_list,IP6_list)
  Begin procedure
    increment  snaptr_output_order;
    create structure with fields
       (hostname, service_and_protocol_set,port,IP4_list,IP6_list,snaptr_output_order,List_Name)
    add structure to end of candidate_list;
    return(looking);
  End procedure;

/*
 * The procedure to find all candidate nodes.
 *
 */

procedure get_candidate_list (targetFQDN, desired_service_and_protocol_set, List_Name)
  Begin procedure
  candidate_list:= empty;
  snaptr_output_order:=0;

  /*
   * The S-NAPTR resolving starts here.
   */
   status:=S_NAPTR_to_callback(targetFQDN,
                               desired_service_and_protocol_set ,
                               private_store_candidate_list);
   return(candidate_list);
  End procedure;
```

The procedure includes the NAPTR output ordering explicitly as a field with each record which is important in the context of "topon" matches and checks for collocated nodes.

C.4 S-NAPTR procedure pseudo-code with topon

Collocation, when stated as specifically applicable in a procedure, takes precedence over other criteria such as topological ordering or S-NAPTR ordering. Topological ordering, when stated as specficially applicable in a procedure, takes precedence over S-NAPTR ordering. However, S-NAPTR ordering is used for ordering nodes with equivalent topological distances. Pseudo code below is informative and shows how to implement the ordering of record selection. See sub-clause 4.3.2 for normative text.

Assume two distinct types of nodes "A" and "B" are being checked for closeness and the best record pair is needed. First step, which is documented in each case in the main text of this document is to get two candidate lists using a procedure such as that outlined in subclause C.3.

```
  candidate_list_A:= get_candidates (targetFQDN_A,desired_service_and_protocol_set_A, "A");
  candidate_list_B:= get_candidates (targetFQDN_B,desired_service_and_protocol_set_B, "B");
```

As an example take the selection of a PGW and SGW by an MME during a UE initial attach procedure. Both a PGW and SGW need to be selected and if "topon" is used in both lists the selected pair is to be as close as possible (collocated being the closest).

Sometimes one list in the procedure is not found by DNS (or was found previously) because the node was already selected. In that case, one of the candidate lists can be just one node.

For example, an UE with an existing PDN connection adds a new PDN connection to a different APN, which may result in a different PGW but needs to continue using the current SGW. Here one of the two candidate lists would just be the data for the current SGW node (i.e., its node name and whether it supports GTPv2 and/or PMIPv6 for S5/S8).

The following pseudo-code illustrates topological matching with full ordering.

```
procedure topo_matching  (candidate_list_A,candidate_list_B, topology_check, colocation_check)//
Comment: topology_check and colocation_check are booleans indicating if those checks are done
  Begin procedure
   paired_sets_list:=empty;
   if (colocation_check)
   Begin if
```

```
     total_list:= candidate_list_A appended with candidate_list B;
// Comment:  Below canonical_node_name is the hostname with first two labels removed
         Foreach unique canonical_node_name from total_list do
           Begin foreach
             Foreach servce and protocol do
               Begin foreach
                full_match_list:= get all records from total_list with service and protocol
                                  and canonical_node name
                If there is at least one "A" record and one "B" record in full_match_list then
                    Begin If
                       degree:=256;
                       suffix:=canonical_node_name;
                       create structure with fields
                         (degree, suffix,service_and_protocol,full_match_list)
                       add that structure to paired_sets_list
                    End If
               End foreach
           End foreach
   End if;
   if (topology_check)         Begin if
           max_labels:= (maximum number of DNS labels in total_list)-  2
         else
           max_labels =0;
         End If;

   from number_labels_to_match:=  max_labels down to 0 do
     Begin do
       if number_labels_to_match equals 0 then
         Begin if
           total_list:= candidate_list_A appended with candidate_list B
         else
           total_list:= get all records from candidate_list_A and candidate_list_B
                        with "topon" as first label and hostname has at least
                        (number_labels_to_match+2) labels
// Comment: Add 2 since the first two labels are not part of the node name         End If;

// Comment:  Below suffix is a hostname chopped off to include only the last number_labels_to_match
         Foreach unique suffix from total_list do
           Begin foreach
             Foreach servce and protocol do
               Begin foreach
                full_match_list:= get all records from total_list with service and protocol
                                  and suffix contained in end of the hostname
                If there is at least one "A" record and one "B" record in full_match_list then
                    Begin If
                       degree:=number_labels
                       create structure with fields
                         (degree, suffix,service_and_protocol,full_match_list)
                       add that structure to paired_sets_list
                    End If
               End foreach
           End foreach
     End do
   sort paired_sets by degree
   return (paired_sets_list)
  End Procedure;
```

NOTE 1: Matching collocated nodes get a degree of 256, which is above any normal match. Also the above procedure is specific to this document and is not a part of S-NAPTR.

NOTE 2: Order from S-NAPTR is from one S-NAPTR procedure. There is no meaningful order obtained from S-NAPTR between records from two different S-NAPTR procedures. So one node type will be selected "logically first" based on other criteria outside S-NAPTR information.

The above procedure simply creates a list of records which are sorted by decreasing degree of matching in DNS labels. It also gives the list of paired nodes with the same suffix and compatible service which is needed by the 3GPP application.

Since highest degree is preferred over S-NAPTR ordering with "topon" labels the selection is done by degree starting with the highest degree obtaining only the possible "A" and "B" nodes at that degree. A sublist of the paired_sets_list containing the highest degree is taken from paired_sets_list denoted as degree_sublist. Assume the "A" node is to be selected "logically first". Sort the "A" parts of degree_sublist in increasing "snaptr_output_order". For the records in that order try to connect to the node with the service and protocol in the record using 3GPP procedures. On failure

proceed with the next record. When that degree_sublist is exhausted then degree-1 is tried and so on until an "A" node is selected. Once an "A" node is chosen, the procedure has also a selected degree, suffix and service.

NOTE 3: The remaining part of this procedure is not needed if the "B" node was already pre-selected outside the present procedure.

Taking the degree_sublist used to select the "A" node create a new sublist from only records with the same service, same protocol and suffix. Remove the "A" node records from that sublist. Sort this new sublist by increasing "snaptr_output_order" (see subclause C.3). Using the records in that order try to connect to the "B" nodes with the service and protocol in the record. On a failure proceed with the next record. When that list is exhausted the procedure continues in next paragraph.

NOTE 4: The suffix of the "A" node that was selected influences which "B" nodes are closest to it. We can't easily and simply reuse the above structure for that reason and it is easier to "reset" the procedure..

A new candidate_list_A is created consisting only of the selected "A" node A and service_and_protocol. The procedure "topo_matching " is run again giving a new paired_sets_list. The "A" node records are removed from the new paired_sets_list leaving only "B" nodes. Sort the records in paired_sets_list in decreasing order of degree and within degree in increasing "snaptr_output_order" (see subclause C.3). Using the records in that order try to connect to the "B" nodes with the service in the record. On failure go to next record.

NOTE 5: Failing to actually contact a node should result in the failing node(s) to be removed from consideration for a period of time. Such removal is not detailed above. Also a reasonable implementation would give up after some maximum number of failed attempts.

Annex D (Informative): Change history

Change history							
Date	**TSG #**	**TSG Doc.**	**CR**	**Rev**	**Subject/Comment**	**Old**	**New**
2008-12	CT#42	CP-080714			V2.0.0 approved in CT#42	2.0.0	8.0.0
2009-03	CT#43	CP-090054	0001	1	DNS examples	8.0.0	8.1.0
			0002	2	DNS cleanup regarding topon/topoff		
			0003	1	Missing sub-clause for node name records of SGW		
			0004	1	DNS Editorial cleanup		
			0005		SGW Selection for TAU and Handover		
			0007	1	Correction to DNS procedures		
			0008	2	Clarifications to NAPTR flag usage		
			0009	1			
2009-06	CT#44	CP-090291	0010	1	Home agent discovery	8.1.0	8.2.0
			0011	1	Correction for PGW selection for colocated PGW/GGSN		
			0012	1	Handling when hostnames are missing topon/topoff		
			0013	1	Release 8 SGSN clarifications		

History

Document history		
V8.0.0	January 2009	Publication
V8.1.0	April 2009	Publication
V8.2.0	June 2009	Publication